

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 20, 2003, 10:40:37 ; Search time 259 Seconds
(without alignments)
6910.144 Million cell updates/sec

Title: US-09-909-207-1

Perfect score: 663

Sequence: 1 CAATGTCACGCAATTC.....TAACCTTGGATAAAACAAT 663

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_15Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	663	100.0	663	17 AAT16101	Xylanase gene. Ba
2	663	100.0	744	17 AAT16102	Xylanase precursor
3	663	100.0	1513	17 AAT16103	Xylanase gene. Ba
4	634.2	95.7	744	19 AAV13067	Glycosyl hydrolase
5	634.2	95.7	871	19 AAV30255	DNA encoding a Bac
6	624.6	94.2	744	16 AAQ92878	Thermotable alkal
7	507.8	76.6	747	19 AAV13074	Xylanase activity
8	255.4	38.5	2364	21 AAZ51821	Clostridium sterco

9	255.2	38.5	1022	16 AAQ80923	B. pumilus xylanase
10	212.2	32.0	1190	18 AAT90972	Nucleotide sequenc
11	204.8	30.9	1244	17 AAT08142	Xylanase xynD gene
12	154.4	23.3	164	16 AAQ92876	Thermotable alkal
13	151.2	22.8	164	16 AAQ92875	Thermotable alkal
14	119.4	18.0	1375	17 AAX90405	Actinomadura sp. D
15	119.4	18.0	1375	18 AAT64930	Actinomadura flexu
16	119	17.9	229	19 AAV15063	Xylanase activity
17	118.4	17.9	573	19 AAV36098	DNA sequence of th
18	117.8	17.8	1207	17 AAT42374	Bacillus sp. therm
19	115.2	17.4	213	19 AAV15059	Xylanase activity
20	115.2	17.4	596	21 AAX48219	T. reesei xylanase
21	115.2	17.4	596	24 AAD29410	Trichoderma reesei
22	111.8	16.9	1273	16 AAQ90388	Xylanase xynA gene
23	108.2	16.3	801	22 AAI66600	Bacillus haloduran
24	106	16.0	2898	21 AAZ46404	Penicillium funicu
25	105.8	16.0	1195	20 AAZ28864	Streptomyces oliva
26	105.8	16.0	1195	20 AAZ28865	Streptomyces oliva
27	104.4	15.7	851	18 AAT63044	Aspergillus niger
28	103	15.5	1281	18 AAT71585	Chaetomium thermop
29	100.8	15.2	1581	15 AAQ69150	Aspergillus tubige
30	100.8	15.2	3105	21 AAZ51819	Ruminococcus xylan
31	100.4	15.1	234	19 AAV15064	Xylanase activity
32	100.4	15.1	807	18 AAT71749	Fibrobacter succin
33	100.4	15.1	3141	18 AAT71746	Fibrobacter succin
34	98.8	14.9	648	18 AAT71748	Fibrobacter succin
35	98	14.8	1027	22 AAF85591	Acrononium cellulo
36	93.4	14.1	675	14 AAQ36563	Endo-xylanase gene
37	92	13.9	164	16 AAQ92877	Thermotable alkal
38	91	13.7	1015	14 AAQ54775	T. reesei xln2 gen
39	91	13.7	1015	20 AAV81332	T. reesei xylanase
40	89.6	13.5	822	19 AAV29598	DNA sequence of th
41	88.2	13.3	234	19 AAV15056	Xylanase activity
42	87.2	13.2	941	14 AAQ54776	T. reesei xln1 gen
43	87.2	13.2	941	20 AAV81331	T. reesei xylanase
44	87.2	13.2	2338	15 AAQ54656	Plasmid pNX1 compr
45	87	13.1	234	19 AAV15058	Xylanase activity

ALIGNMENTS

RESULT 1

AAT16101
ID AAT16101 standard; DNA; 663 BP.

XX AAT16101;

DT 15-MAY-1996 (first entry)

XX Xylanase gene.

XX Xylanase; thermotable enzyme; paper; pulp; bleaching; feedstuff; baking; 68.

OS Bacillus sp. strain 710/1 (LMG P-14798).

XX Key Location/Qualifiers

PH mat_peptide 1..663

FT /**tag= a

FT /EC_number= 3.2.1.8

XX AU9525086-A.

XX 08-FEB-1996.

XX 19-JUL-1995; 95AU-0025086.

XX 17-MAY-1995; 95BE-0000448.

XX 26-JUL-1994; 94BE-0000706.

XX (SOLV) SOLVAY SA.

PA (SOLV) SOLVAY & Cie.

```

XX De Buyl E, Detroz R, Lahaye A, Ledoux P;
XX WPI; 1996-117341/13.
XX P-PSDB; AAR92053.
XX Bacillus derived xylanase active over wide pH range - used in
XX treatment of paper pulp, animal feeds and in bakery goods
XX
XX Claim 30; Page 50-51; 94pp; English.
XX
XX A DNA sequence (AAT16101) coding for a thermostable mature xylanase
XX (AAR92053) was isolated from a gene library of Bacillus sp. 720/1
XX (LMG P-14798). Sequences were also obtd. (see AAT16102 and AAT16103)
XX for the xylanase precursor and for the complete gene including 5'
XX and 3' untranslated sequences. DNA coding for the mature enzyme
XX may be incorporated into a vector and expressed from either its
XX own promoter or from the Bacillus pumilus PRL B12 promoter (AA073996),
XX and used for prodn. of recombinant xylanase in transformed hosts,
XX pref. Bacillus licheniformis or B. pumilus. The enzyme is useful
XX in the paper-pulp, animal feed and baking industries.
XX
XX Sequence 663 BP; 220 A; 131 C; 146 G; 166 T; 0 other;
XX
XX Query Match 100.0%; Score 663; DB 17; Length 663;
XX Best Local Similarity 100.0%; Pred. No. 1.3e-188;
XX Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CAATCGTCACGCAATTCATTTGGCAACGAGTGGCTATGATTAATTTGGAAA 60
DB 1 CAATCGTCACGCAATTCATTTGGCAACGAGTGGCTATGATTAATTTGGAAA 60
QY 61 GATAGCGGTGGCTCTGGACAATGATTCATTCATGAGCGGTACGTTCCAGTCCCAATGG 120
DB 61 GATAGCGGTGGCTCTGGACAATGATTCATTCATGAGCGGTACGTTCCAGTCCCAATGG 120
QY 121 AACATGTTAACACATATTTATTCCTTAAGGTAAGTAATTCATTAACACAAACACAC 180
DB 121 AACATGTTAACACATATTTATTCCTTAAGGTAAGTAATTCATTAACACAAACACAC 180
QY 181 CAACAAGTGTGAACATGTCATTAACCTACGAGCCCACTTCACCAACCAATGGTAATGG 240
DB 181 CAACAAGTGTGAACATGTCATTAACCTACGAGCCCACTTCACCAACCAATGGTAATGG 240
QY 241 TATTTATGCGTCTATGATGTTGACCTCTTTCGTAATTAATTTGTCGACAGT 300
DB 241 TATTTATGCGTCTATGATGTTGACCTCTTTCGTAATTAATTTGTCGACAGT 300
QY 301 TGGGGCAACTGGCGTCCACGAGAGCAAGCCCTAAGGGAGCAATCTGTTGATGGAGA 360
DB 301 TGGGGCAACTGGCGTCCACGAGAGCAAGCCCTAAGGGAGCAATCTGTTGATGGAGA 360
QY 361 ACATATGATATCTACGAGACTCTTAAGATCAATCAACCTCCATTAAAGGGATGGCACA 420
DB 361 ACATATGATATCTACGAGACTCTTAAGATCAATCAACCTCCATTAAAGGGATGGCACA 420
QY 421 TTTAAACAATATTGAGGTTCGAAGATCGAAACGACGAGTGGACGATTTCTGTACG 480
DB 421 TTTAAACAATATTGAGGTTCGAAGATCGAAACGACGAGTGGACGATTTCTGTACG 480
QY 481 AACCACTTAGAGCGTGGGAAAACTTAGAGATCAATTAAGGGAAAAATGTAAGTCGGG 540
DB 481 AACCACTTAGAGCGTGGGAAAACTTAGAGATCAATTAAGGGAAAAATGTAAGTCGGG 540
QY 541 CTTACTGTGAAGGCTATCAAAAGTAGCGGAAGTCTAATGTAATGACATACATPAGA 600
DB 541 CTTACTGTGAAGGCTATCAAAAGTAGCGGAAGTCTAATGTAATGACATACATPAGA 600
QY 601 ATTTAAGGTACCCCTCTCTCAACTATTAATGATGAGAGACATACTTTGATATAAAC 660
DB 601 ATTTAAGGTACCCCTCTCTCAACTATTAATGATGAGAGACATACTTTGATATAAAC 660
QY 661 AAT 663

```

```

DB 661 AAT 663
|||
RESULT 2
ID AAT16102 standard; DNA; 744 BP.
AAT16102;
15-MAY-1996 (first entry)
Xylanase precursor gene.
Xylanase; thermostable enzyme; paper; pulp; bleaching; feedstuff;
baking; ss.
Bacillus sp. strain 710/1 (LMG P-14798).
Key Location/Qualifiers
sig_peptide 1..81
mat_peptide 82..744
/*tag= b
/EC_number= 3.2.1.8
AU9525086-A.
08-FEB-1996.
19-JUL-1995; 95ANU-0025086.
17-MAY-1995; 95BE-0000448.
26-JUL-1994; 94BE-0000706.
(SOLV ) SOLVAY SA.
(SOLV ) SOLVAY & CIE.
De Buyl E, Detroz R, Lahaye A, Ledoux P;
WPI; 1996-117341/13.
P-PSDB; AAR92054.
Bacillus derived xylanase active over wide pH range - used in
treatment of paper pulp, animal feeds and in bakery goods
Claim 9; Page 54-55; 94pp; English.
XX
XX A DNA sequence (AAT16102) coding for a xylanase precursor (AAR92054)
XX was isolated from a gene library of Bacillus sp. 720/1 (LMG P-14798).
XX The sequence (AAT16103) for the complete gene including 5' and 3'
XX untranslated sequences was also obtd. The gene may be incorporated
XX into a vector and expressed from either its own promoter or from the
XX Bacillus pumilus PRL B12 promoter (see AA073996), and used for prodn. of
XX recombinant thermostable xylanase in transformed hosts, pref. Bacillus
XX licheniformis or B. pumilus. The enzyme is useful in the paper-pulp,
XX animal feed and baking industries.
XX
XX Sequence 744 BP; 246 A; 144 C; 160 G; 194 T; 0 other;
XX
XX Query Match 100.0%; Score 663; DB 17; Length 744;
XX Best Local Similarity 100.0%; Pred. No. 1.4e-188;
XX Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 CAATCGTCACGCAATTCATTTGGCAACGAGTGGCTATGATTAATTTGGAAA 60
DB 82 CAATCGTCACGCAATTCATTTGGCAACGAGTGGCTATGATTAATTTGGAAA 141
QY 61 GATAGCGGTGGCTCTGGACAATGATTCATTCATGAGCGGTACGTTCCAGTCCCAATGG 120
DB 142 GATAGCGGTGGCTCTGGACAATGATTCATTCATGAGCGGTACGTTCCAGTCCCAATGG 201
QY 121 AACATGTTAACACATATTTATTCCTTAAGGTAAGTAATTCATTAACACAAACACAC 180

```



```
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI
XX PI Outtrup H, Biegard-frantzen H, Schuelein M, Olsen AA, Jorgensen PL,
XX PI Dambmann C;
XX DR WPI; 1998-376805/32.
XX DR P-PSDB; AAW60562.
XX
XX DNA construct encoding Bacillus agaradherens xylanolytic enzyme -
XX PT and vectors and Bacillus cells containing these, useful for
XX PT recombinant production of the enzyme for use in agricultural waste
XX PT breakdown and lignocellulosic material treatment
XX
XX PS Claim 1; Columns 11-14; 10pp; English.
XX
XX The present sequence encodes a xylanolytic enzyme of Bacillus
XX CC agaradherens NCIMB 40482. Xylanolytic enzymes are used for enzymatic
XX CC breakdown of agricultural wastes for production of alcohol fuels,
XX CC enzymatic treatment of animal feeds to release free pentose sugars,
XX CC manufacturing of dissolving pulps yielding cellulose and bio-bleaching
XX CC of wood pulp. They are also used for treatment of lignocellulosic
XX CC material e.g. paper and pulp, or as an animal feed additive.
XX CC (Updated on 25-MAR-2003 to correct PR field.)
XX CC (Updated on 25-MAR-2003 to correct PI field.)
XX
XX SQ Sequence 871 BP; 283 A; 171 C; 179 G; 238 T; 0 other;
XX
Query Match 95.7%; Score 634.2; DB 19; Length 871;
Best Local Similarity 97.3%; Pred. No. 6.5e-180;
Matches 645; Conservative 0; Mismatches 18; Indels 0; Gaps 0;
QY 1 CAATCGTCACCGACAAATTCATTCGGCAACACGATGGCTATGATTATGAATTTGGAAA 60
DB 82 CAATCGTCACCGACAAATTCATTCGGCAACACGATGGCTATGATTATGAATTTGGAAA 141
QY 61 GATACGGTGGCTCTGGGCAATGATTTCTCAATCATGGCGGTACGTTCAAGTCCCAATGG 120
DB 142 GATACGGTGGCTCTGGGCAATGATTTCTCAATCATGGCGGTACGTTCAAGTCCCAATGG 201
QY 121 AACATGTTTAAACAATATTTCCGTAAAGTAAATAATTCATGAACACACACACAC 180
DB 202 AACATGTTTAAACAATATTTCCGTAAAGTAAATAATTCATGAACACACACACAC 261
QY 181 CAACAAGTTGTAACATGTCCTCAATAAATACGAGGCAATCTTCCAAACCAATGGTAATCGG 240
DB 262 CAACAAGTTGTAACATGTCCTCAATAAATACGAGGCAATCTTCCAAACCAATGGTAATCGG 321
QY 241 TATTTATCGGCTATGTTGGACTGTTGACCCCTTTGTCGAATATTAATTTGTCGACAGT 300
DB 322 TATTTATCGGCTATGTTGGACTGTTGACCCCTTTGTCGAATATTAATTTGTCGATAGT 381
QY 301 TGGGCACTGGCGTCCACGAGGACGACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360
DB 382 TGGGCACTGGCGTCCACGAGGACGACGCTTAAGGGGACCATCACTGTTGATGGAGGA 441
QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCAATTAAGGGATGCCACA 420
DB 442 ACATATGATATCTAGAACTCTTAGAGTCAATCAAGCCCTCAATTAAGGGATGCCACA 501
QY 421 TTTAAACAATATTTGAGTGTTCGAAGATCGAAACCGCAGGTGGCAATTTCTGTACG 480
DB 502 TTTAAACAATATTTGAGTGTTCGAAGATCGAAACCGCAGGTGGCAATTTCTGTACG 561
QY 481 AACCACTTTAGCGTGGGAAACTTAGGGATGATATGGGAAATGTATGAAGTCCGG 540
DB 562 AACCACTTTAGCGTGGGAAACTTAGGGATGATATGGGAAATGTATGAAGTCCGG 621
QY 541 CTTACTGTAGAGGCTATCAAGTAGTGGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600
DB 622 CTTACTGTAGAGGCTATCAAGTAGTGGGAAGTGTCTAATGTATATAGCAATACACTAAGA 681
QY 601 ATTAACGGTAACCCCTCTCTCAACTATTAGTAATATGACGAGACATAACTTTTGGATAAAAAC 660
DB
XX
XX Db 682 ATTAACGGTAACCCCTCTCTCAACTATTAGTAATGACAGACATAAATCTAGATAAAAAC 741
XX QY 661 AAT 663
XX Db 742 AAT 744
XX
XX RESULT 6
XX AAQ92878
XX ID AAQ92878 standard; DNA; 744 BP.
XX AC AAQ92878;
XX
XX DT 12-FEB-1996 (first entry)
XX
XX DE Thermostable alkaline endo-1,4-beta-D-xylanase gene.
XX
XX KW thermostable alkaline endo-1,4-beta-D-xylanase gene; cloning;
XX KW polymerase chain reaction; Escherichia coli; EC-3.2.1.8; paper;
XX KW pulp; bleaching; ds.
XX
XX OS Bacillus sp. 1-43-3 (CBS 672.93).
XX
XX PN WO9518219-A1.
XX
XX PD 06-JUL-1995.
XX
XX PF 23-DEC-1994; 94WO-EP04312.
XX
XX PR 24-DEC-1993; 93EP-0203694.
XX
XX PA (KONN ) GIST-BROCADES NV.
XX
XX PI Farrell RL, Goedegebuur F, Herbes WT, Herweijer MA;
XX PI Iverson S, Jones BE, Quax WJ, Van Beckhoven RFWC;
XX PI Van der Kleij WA, Van Solingen P, Williams DP;
XX
XX DR WPI; 1995-246385/32.
XX P-PSDB; AAR76551.
XX
XX Novel xylanase enzyme active at high pH - useful in paper and pulp
XX prodn. processes
XX
XX Claim 4; Page 42-43; 54pp; English.
XX
XX The sequence encodes an alkaline endo-1,4-beta-D-xylanase (G-type)
XX from Bacillus sp. 1-43-3. Fragments of the sequence (e.g. the
XX internal fragments given in AAQ92876 and AAQ92877) may be amplified
XX by polymerase chain reaction, e.g. using primers with sequences
XX AAQ92866, AAQ92867, AAQ92868 and AAQ92869. The DNA may be cloned in
XX Escherichia coli using a plasmid vector for recombinant xylanase
XX production. The xylanase may be used in the paper and pulp
XX industries, where it produces an increase in % ISO brightness of
XX softwood pulp of at least 1.0 over non-enzymatically treated pulp
XX in an ECF pulp bleaching process, at pH 9.0 and 65 deg C. The
XX enzyme may be used in production of paper, board and fluff pulp,
XX and has low cellulase activity. The increased brightness
XX produced using the xylanase allows reduction in the amount of
XX bleaching chemicals used.
XX
XX SQ Sequence 744 BP; 248 A; 145 C; 159 G; 192 T; 0 other;
XX
Query Match 94.2%; Score 624.6; DB 16; Length 744;
Best Local Similarity 96.4%; Pred. No. 4.6e-177;
Matches 639; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 1 CAATCGTCACCGACAAATTCATTCGGCAACCAACGATGGCTATGATTATGAATTTGGAAA 60
DB 79 CAATCGTCACCGACAAATTCATTCGGCAACCAACGATGGCTATGATTATGAATTTGGAAA 138
QY 61 GATACGGTGGCTCTGGGACAAATGATTCTCAATCATGGCGGTACGTTCAAGTCCCAATGG 120
DB
```


Db 682 ATTAACGGTACCCCTCTCTCAACTATTAGTAATGACAAGAGCATAACTCTAGATAAAAC 741
 QY 661 AAT 663
 ||||
 Db 742 AAT 744

RESULT 8

AAZ51821
 ID AAZ51821 standard; DNA; 2364 BP.

XX.

AC AAZ51821;

XX

DT 04-JUL-2000 (first entry)

XX

DE Clostridium stercorarium xylanase A DNA.

XX Xylanase A; XynA; cellulosome; feruloyl esterase; phenolic acid esterase;
 KW thermostable; ferulic acid; wheat bran; agricultural byproduct; treat;
 KW grass; paper and pulp industry; feed processing; food additive;
 KW plant cell wall material; degradation; ds.

XX Clostridium stercorarium.

OS

XX

FH Key Location/Qualifiers

FT CDS 440..11978

FT /tag= e

FT /product= "Xylanase A"

XX

XX

PN WO200014243-A1.

XX

PD 16-MAR-2000.

XX

XX

PF 03-SEP-1999; 99WO-US20304.

XX

XX

PR 04-SEP-1998; 98US-0099136.

XX

XX

PA (UTGE-) UNIV GEORGIA RES FOUND INC.

XX

XX

PI Blum DL, Kataeva I, Li X, Ljungdahl LG;

XX

XX

DR WPI; 2000-256991/22.

XX

DR P-P8DB; AAY70523.

XX

XX

PT New recombinant DNA molecule comprising a sequence encoding feruloyl

XX

PT esterase protein, useful for treating grasses and other plant materials

XX

PT used in pulp and paper industries, feed processing and food additives -

XX

XX

PS Disclosure; Pages 97-99; 105pp; English.

XX

XX

CC The present sequence is a DNA encoding xylanase A (XynA) from

XX

CC Clostridium stercorarium. XynA contains family VI cellulose

XX

CC binding domain (CBD) which is homologous to the CBD of xylanase Z

XX

CC (XynZ) of Clostridium thermocellum. XynZ is an enzymatic component of

XX

CC C. thermocellum cellulosome and has a multi-domain structure which

XX

CC includes a dockerin domain, a catalytic xylanase domain,

XX

CC a family VI cellulose binding domain and a domain of

XX

CC unknown function. The unknown domain in the N-terminal region of XynZ

XX

CC has been found to contain feruloyl (phenolic acid) esterase which

XX

CC is involved in the degradation of plant cell wall material. The novel

XX

CC feruloyl esterase is thermostable, easy to purify, has high temperature

XX

CC optima and stable over a wide pH range. The enzyme is used for

XX

CC producing ferulic acid from wheat bran or agricultural byproducts,

533 CGAATAATTTACGACCAATGACAGGACACATCGAGGCTACGACTATGAGCTCTGGAAA 592
 QY 61 GATAGCGTGGCTCTGGGACAAATGATCTCAATCATGCGCGTACGTTCCAGTGCCTCAATGG 120
 Db 593 GA---CTACGGGAATACGATTTATGGAATTAACGACGGTGGTACTTTTAGTTGTCAATGG 649
 QY 121 AACAAATGTTAAACAATATTAATTCGTTAAAGGTAAAAAATCAATGAACAACAACACAC 180
 Db 650 AGTAATATCGGTAATGCACATATTTAGAAAAGGGAGAAAATTTAATTCGACAAAACCTAT 709
 QY 181 CAACAAGTTGGTAACATGTCCATAAATACGAGGACCACTTCCAAACCAAAATGGTATGG 240
 Db 710 CAAGAAATTTAGGAGACATAGTAGTTGAATATGCTGTGTGATTACAATCCAAACCGAAATTC 769
 QY 241 TATTTATGCGTCTATGTTGGACTGTTGACCTCTTGTGCAATATTAATTTGTCGACAGT 300
 Db 770 TATTTGTTGTTTACGGTTGGACAAAGAAATCCACTGTTGAATATTACATTTGTAGAAAGC 829
 QY 301 TGGGGCAACTCGGCTCCACAGGAGCAACGCTTAAGGGGACCATCAC---TGTGTGATGGA 357
 Db 830 TGGGGCAGCTGGCGTCCACCTGGAGCAACACCCAAAGGAACCATCACAGTGGATGGCA 889
 QY 358 GGAACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTCGC 417
 Db 890 GGTACTTATGAAATATATGAAACTACCGGGTAAATCAGCCTTCATCGATGGAACTGG 949
 QY 418 ACATTTAAACAATATTTGGAGTGTTCGAAAGATCGAAACGACGAGTGGGACCAAGTTCTGTC 477
 Db 950 ACATTTCCACAAATATTTGGAGTGTTCGATCATCCAGAGAACACGGGAACAATATCTGTC 1009
 QY 478 AGCAACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATGGGAAAATGTATGAAGTC 537
 Db 1010 ACTGAACATTTTAAACAGTGGGAAAGAAATGGGCATCGAATGGGTAGATGTATGAAGTT 1069
 QY 538 CGGCTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTATGTATATAGCAATACACTA 597
 Db 1070 GCTCTTACCGTTGAAGGTTATCAGAGCAGTGGGTACGCTAATGTATACAGAATGAATC 1129
 QY 598 AGAATTAACGGTAACCC 614
 Db 1130 AGAATAGTGCAAAATCC 1146
 RESULT 9
 AAQ80923
 ID AAQ80923 standard; DNA; 1022 BP.
 XX
 AC AAQ80923;
 XX
 DT 02-AUG-1995 (first entry)
 DE B. pumilus xylanase gene.
 KW Xylanase; endo-1,4-beta-D-xylanase; wood pulp; pulping;
 KW biobleaching; bleaching; Bacillus licheniformis; ss.
 XX Bacillus pumilus.
 OS
 FH Key Location/Qualifiers
 FT 5'UTR 1..185
 FT /tag= a
 FT /note= "Claim 12"
 FT misc_difference 107
 FT /tag= b
 FT /note= "Base n at position 107 is not identified
 in the specification"
 FT CDS 186..869
 FT /tag= c
 FT sig_peptide 186..266
 FT /tag= d
 FT /note= "Claim 11"
 FT mat_peptide 267..866

Query Match 38.5%; Score 255.4; DB 21; Length 2364;
 Best Local Similarity 65.6%; Pred. No. 4.2e-66;
 Matches 405; Conservative 0; Mismatches 206; Indels 6; Gaps 2;
 QY 1 CAATCGTCACCGACAATTCATTTGGCAACCAACGATGGCTATGATTAATTTGGAAA 60

```
/*tag= e
/EC_number= 3.2.1.8
/note= "claim 10"
```

GB2279955-A.

18-JAN-1995

15-JUL-1993; 93GB-0014780.

15-JUL-1993; 93GB-0014780.

(SOLV) SOLVAY SA.
(SOLV) SOLVAY & CIE.

Amory A, Andre C, De Buyl E, Detroz R, Lahaye A,
Ledoux P,

WPI; 1995-039214/06.
P-PSDB; AAR68849.

Claim 13; Fig. 1a-1b; 97pp; English.

Claim 13; Fig. 1a-1b; 97pp; English

A *Bacillus pumilus* PTL B12 (ATCC 55443) gene library was screened for recombinant plasmids carrying the xylanase gene. A chromosomal fragment obtained from isolate PBKPI was subcloned and expressed in *Escherichia coli* JM109. The sequence of the xylanase gene (as a 1022 bp Sau3AI fragment) carried by a selected transformant is given in AA080923.

SQ Sequence 1022 BP; 329 A; 197 C; 229 G; 266 T; 1 other,

Query match 38.5%; Score 255.2; DB 16; Length 1022;

Best Local Similarity 65.5%; Pred. NO. 3.4e-66;
Matches 406; Conservative 0; Mismatches 208; Indels 6; Gaps 2

Oy	13	GACATATTCATTTGGGCAACGACGATGGCTTGTATGATTAATGATTTTGGAAAGATACGGTGGC	72
Db	279	GATATATAGATTAGGGAACACACAGCGGATACGATTTTGGATTATAGAAAGATTAC---GGA	335
Oy	73	TCCTGGACAATGATTTCTCAATCATGCGGTCAGTTCAGTGGCCCAATGAAACAATGTTAAC	132
Db	336	AATACCTCGATGACACTCAATTAACGGCGGGGCTTTAGTGCACATGGAACAATATATTGGA	395
Oy	133	AACATATTTATTCGGTAAGAAGTAAAAATTCATGAAACAACAACACCAACAAGTTGCT	192
Db	396	AATGCGTTATTTTGGAAAGAAAGAAAGTTGATTCACCTAAATCATATCACTCAACTTGGC	455
Oy	193	AACATGTCCATNAACTACGAGGCCAACTTCCACCAATGATATGCTATTTATGCGTC	252
Db	456	AACATCTCCATCACTACAAACGACGCTTTAACCGGGGGGAAATTCATTTATATGTC	515
Oy	253	TATGCTTGACGTTGGACCTCTTGTGCAATATTATTTGTGCACAGTTGGGGCAACTGG	312
Db	516	TATGCTGTGACCAATCTTCATTTAGCTGAAATCTACATTTGTTAGTCACTGGGGCAACATAT	575
Oy	313	CGTCACCAAGAGCAACGCCCTAAGGGGACATCACCTGTTGATGAGAGAAACATATGATATC	372
Db	576	CGTCCAAAGG---AACGTATTAAGGATCATTTTATGCGGATGGAGGACATATGACATA	632
Oy	373	TACGAGACTTTAGATCAATCAACCCCTCATTTAAGGGGATTCGCATTTAAACAATAT	432
Db	633	TATGAAACGCTCCGATCATCAAGCCTTCTATCATTTGGAGACGCTACCTTCAACAATAT	692
Oy	433	TGGAATGTTCCGAAGTCGAACGCCAGATGGGACGATTTCTTCAGCAACCACTTTAGA	492
Db	693	TGGAATGTATGCTCAACCAACGCAACGGAACGCTCTCCGTCAGTGGCAATTTTAA	752

Qy	493	GGCTGGGAAAACCTTGGGATGTAATGGGAAAATGTGTGAAGTCGGCTTACGTGTAA	555
Db	753	AAATGGGAAAGCTTGGCATGCCAATGGGAAAAATGTTATGAACAGCATTTACGTGTAA	812
Qy	553	GGCTATCAAAAGTAGCGGAAGTCTAATGTAATATAGCATTAACATAAGAAATTAACGTATAC	612
Db	813	GGGTACGGAAGCAAGGAAGTCGAATGATCATAGCAATCAAGCTGATGATTCGATTAABAG	872
Qy	613	CCCTCTCTCAACTTATAGTAA	632
Db	873	CATATGAAAAAGCCAGCAA	892

RESULT 10

ID AAT90972 standard; DNA; 1190 BP

AC AAT90972;

DT 22-MAY-1998 (first entry)

DE Nucleotide sequence encoding an enzyme with xylanase activity.

KM xylnb gene; xylanase; enzyme; *Dicoryoglossum thermophilum* strain Rt46B.1.2
KM G-Xylanase; beta -1,4-xylanase activity; activity; thermal stability; ss
KM biological bleaching; cellulose product; paper pulp; ss.

Dictyoglomus thermophilum.

FH	Key	Location/Qualifiers
FM	000	07 1140

ET

FT	/product= xylinase
ET	67 138

3
6
7
8
9
10
11
12

FT
yy

PN W09736995-A2
XX

PD 09-OCT-199

PF 27-MAR-1997; 97WO-NZ00042
XX

PR 29-MAR-1996; 96NZ-0286296
XX

PA (PACI-) PACIFIC ENZYMES LTD
XX

XX
P1 Bergquist PL, Gibbs MD, Morris D,

DR WP1; 199/-503090/46
DB P-PSDB: AAW30267

Dictyodolum thermophilum

PT thermally stable enzymes for the bleaching of cellulose products especially paper pulp

XX Disclosure: Fig 3. 1

xx The present sequence

CC xylanase-enzyme. The present sequence is derived from a gene contained
CC within *Diclostoglossum thermophilum* strain R46B.1. The xylanase is
CC contained within the family of enzymes known as G-Xylanases, and has
CC beta-1,4-xylanase activity. The enzyme has high activity and high
CC thermal stability with optimum activity at 85 degrees Celsius and
CC pH 6.5. The xylanase enzyme is used for the biological bleaching of
CC cellulose products, especially paper pulp. Use of the enzyme ensures
CC that waste streams from the biological bleaching will include less
CC toxic material.

SQ Sequence 1190 BP; 386 A; 197 C; 277 G; 330 T; 0 other;

Query Match 32.0%; Score 212.2; DB 18; Length 1190;

Best Local Similarity 62.4%; Pred. No. 2.9e-53;
Matches 369; Conservative 0; Mismatches 213; Indels 9; Gaps 2;

Qy 3 AATCGTCACGCAAAATTCCATTGGCAACACAGCATGGCTATGATTAATTTTGGAAAAGA 62
Db |||||
150 AACACTAACGAATAAATGCAAGCGGTACTTTTGATGGCTACTACTATGAACATATGGAAGA 209
Qy 63 TAGCGGTGCTTCGGGACAATGATTTCTCAATCATGCCGTCAGTTTCAGTGCCTCAATGAA 122
Db |||||
210 TA---CAGGGAATACAAACCATGACTATATACACACAGGAAGTTTTAGCTGTGTCAGTGGAG 266
Qy 123 CAATGTTTACACATATTTTCGTTAAAGTAAATTTCAATGAAGAAACACAAACACACACA 182
Db |||||
267 CAATATAAACAATGCAATTAATTCAGAACAGGTGAAGTACAAAC-----CAAACCTGGCA 320
Qy 183 ACAAGTTGGTAAACATGTCATATAAACTTACGGAGCCAACTTCCAACCAAAATGGTAATGCGTA 242
Db |||||
321 GTCAATTAGGCACTATTAGAATCACCTACTCAGCCACATATAATCTTAATGGTAATCTCTA 380
Qy 243 TTTATCGGCTATAGTTGGTGAAGTCTTGAACCTCTTGTGCAATATTAATTTGTCGACAGTTG 302
Db |||||
381 CTTATGATCTATGTTGGTGGTCTACTAACTCTTTAGTAGAGTTTTACATTGTAGAAAGTTG 440
Qy 303 GGCACTCGGCTCCACGAGGACAGCAAGCTTAAGGGGACCATCACTGTTGATGGAGAAC 362
Db |||||
441 GGGTAATGGCGTCCACCAAGGTGCAACCTCTCTTGACACAGGTACTATCGACGGTGGTAC 500
Qy 363 ATATGATATCTACGACACTCTTAGAGTCAATCAACCTCCATTAAGGGGATGGCCACATT 422
Db |||||
501 CTATGACATTTACAGAACTACCGGTGTAATCAGCCATCTATTTGTCGGTACAGCTACTTT 560
Qy 423 TAAACAATATGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGGATTTCTGTGACGAA 482
Db |||||
561 TGATCAATATGGAGGTGAAGACATCTAAGAGAACAGTGGAAACAGTCACTGTGAACAGA 620
Qy 483 CACATTAGAGGTGGGAAAACTTACGGATGAATATGGGAAAAATGTATGAAGTGGCGCT 542
Db |||||
621 TCACATTAGGGCATGGGCAATAGAGGTTTAAACCTTGGTACTATTTGATCAAAATTAATCT 680
Qy 543 TACTGTAGAAGGCTAACAAGTAGCGGAAGTCTCAATGCTATATACCAATAC 593
Db |||||
681 TTGTGTTGAAGGATATCAAGCAGTGTTTCGGCTAATATAACACAAAAATAC 731

RESULT 11
AAT08142
ID AAT08142 standard; DNA; 1244 BP.
XX AC AAT08142;
XX AC
DT XX
DT XX
DT 25-MAR-2003 (updated)
DT 12-MAY-1996 (first entry)
XX Xylanase xynd gene.
XX DE
XX XX
XX Xylanase D; xynd gene; thermostable enzyme; paper; pulp; lignin;
KW KW delignification; xylan; bleaching; ds.
OS Extremophile isolate TG456 (CBS 213.94).
XX Key Location/Qualifiers
FH 1..1107
FT /*tag= a
XX .
XX W09534662-A1.
PN XX
XX XX
PD 21-DEC-1995.
PP XX
PF 14-JUN-1995; 95WO-BP02299.
PX XX
PR 14-JUN-1994; 94EP-0201699.
PA (KONN) GIST-BROCADES BV.

RESULT 12
AA092876
ID AA092876 standard; DNA; 164 BP.
XX
XX
AC AA092876;
XX
XX
DT 19-APR-1996 (first entry)
XX
XX
DE Thermostable alkaline xylanase gene internal fragment.
XX
XX
KM Thermostable alkaline endo-1,4-beta-D-xylanase gene; cloning;
XX
XX
KW polymerase chain reaction; Escherichia coli; EC-3.2.1.8; paper;
XX
XX
KW pulp; bleaching; ds.
XX
XX
OS Bacillus sp. 1-43-3 (CBS672.93).
XX
XX
PN WO9518219-A1.
XX
XX
PD 06-JUL-1995.
XX
XX
PF 23-DEC-1994; 94MO-EP04312.
XX
XX
PR 24-DEC-1993; 93EP-0203694.
XX
XX
PA (KONN) GIST-BROCADES NV.
XX
XX
PI Farrell RL, Goedegebuur F, Herbes WT, Herweijer MA;
XX
XX
PI Iverson S, Jones BE, Quax WJ, Van Beekhoven RFMC;
XX
XX
PI Van der Kleij WA, Van Solingen P, Williams DP;
XX
XX
DR WPI; 1995-246385/32.
XX
XX
PT Novel xylanase enzyme active at high pH - useful in paper and pulp
XX
XX
PT prodn. processes
XX
XX
PS Example 7; Page 40-41; 54pp; English.
XX
XX
XX AA092871-77 are examples of several different internal xylanase gene
XX
XX
CC fragments which were used as specific probes to isolate full length
XX
XX
CC cloned genes. AA092876 is from an G1-type xylanase. The DNA may be
XX
XX
CC cloned in E. coli using a plasmid vector, and e.g. clone KEX106 was
XX
XX
CC isolated, producing 23.7 U/ml recombinant xylanase. The xylanase may be
XX
XX
CC used in the paper and pulp industries, where it produces an increase in
XX
XX
CC ISO brightness of softwood pulp of at least 1.0 over non-enzymatically
XX
XX
CC treated pulp in an ECF pulp bleaching process, at pH 9.0 and 65 deg
XX
XX
CC C. The enzyme may be used in production of paper, board and fluff
XX
XX
CC pulp, and has low cellulase activity. The increased brightness
XX
XX
CC produced using the xylanase allows reduction in the amount of
XX
XX
CC bleaching chemicals used.
XX
XX
SQ Sequence 164 BP; 50 A; 35 C; 39 G; 40 T; 0 other;
XX
XX
Query Match 23.3%; Score 154.4; DB 16; Length 164;
Best Local Similarity 96.3%; Pred. No. 2.7e-36;
Matches 158; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

XX
XX
AC AA092875;
XX
XX
DT 19-APR-1996 (first entry)
XX
XX
DE Thermostable alkaline xylanase gene internal fragment.
XX
XX
KM Thermostable alkaline endo-1,4-beta-D-xylanase gene; cloning;
XX
XX
KW polymerase chain reaction; Escherichia coli; EC-3.2.1.8; paper;
XX
XX
KW pulp; bleaching; ds.
XX
XX
OS Bacillus sp. 1-25-2 (CBS671.93).
XX
XX
PN WO9518219-A1.
XX
XX
PD 06-JUL-1995.
XX
XX
PF 23-DEC-1994; 94MO-EP04312.
XX
XX
PR 24-DEC-1993; 93EP-0203694.
XX
XX
PA (KONN) GIST-BROCADES NV.
XX
XX
PI Farrell RL, Goedegebuur F, Herbes WT, Herweijer MA;
XX
XX
PI Iverson S, Jones BE, Quax WJ, Van Beekhoven RFMC;
XX
XX
PI Van der Kleij WA, Van Solingen P, Williams DP;
XX
XX
DR WPI; 1995-246385/32.
XX
XX
PT Novel xylanase enzyme active at high pH - useful in paper and pulp
XX
XX
PT prodn. processes
XX
XX
PS Example 7; Page 40; 54pp; English.
XX
XX
XX AA092871-77 are examples of several different internal xylanase gene
XX
XX
CC fragments which were used as specific probes to isolate full length
XX
XX
CC cloned genes. AA092875 is from an G-type xylanase. The DNA may be cloned
XX
XX
CC in E. coli using a plasmid vector, and e.g. clone KEX106 was isolated,
XX
XX
CC producing 23.7 U/ml recombinant xylanase. The xylanase may be used in
XX
XX
CC the paper and pulp industries, where it produces an increase in ISO
XX
XX
CC brightness of softwood pulp of at least 1.0 over non-enzymatically
XX
XX
CC treated pulp in an ECF pulp bleaching process, at pH 9.0 and 65 deg
XX
XX
CC C. The enzyme may be used in production of paper, board and fluff
XX
XX
CC pulp, and has low cellulase activity. The increased brightness
XX
XX
CC produced using the xylanase allows reduction in the amount of
XX
XX
CC bleaching chemicals used.
XX
XX
SQ Sequence 164 BP; 50 A; 35 C; 39 G; 40 T; 0 other;
XX
XX
Query Match 22.8%; Score 151.2; DB 16; Length 164;
Best Local Similarity 95.1%; Pred. No. 2.5e-35;
Matches 156; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

XX
XX
AC AA092875;
XX
XX
DT 25-MAR-2003 (updated)

DT 28-SEP-1999 (first entry)
 XX Actinomadura sp. DSM43186 35 kDa xylanase encoding DNA.
 DE
 XX
 KW Actinomadura sp. DSM43186; xylanase; Actinomadura flexuosa; lignin;
 KW thermostable; biobleaching; wood pulp; bleaching; hemicellulase;
 KW paper processing; hemicellulose; ss.
 XX
 XX Actinomadura flexuosa.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 303..1337
 FT /*tag= a
 FT /product= "xylanase"
 XX
 XX US5935836-A.
 XX
 XX 10-AUG-1999.
 XX
 XX 06-JUN-1995; 95US-0468812.
 XX
 XX 06-JUN-1995; 95US-0468812.
 PR 29-JUL-1994; 94US-0282001.
 PR 31-OCT-1994; 94US-0332412.
 PR 31-JUL-1995; 95FI-0003639.
 XX
 XX (ALKO-) ALKO GROUP LTD.
 PA (ALKO-) ALKO-YHTIOET OY.
 PA (ROHM) ROEHM ENZYME FINLAND OY.
 XX
 PI Fagerstrom R, Kristo P, Lahtinen T, Lantto R, Maentylae A;
 PI Paloheimo M, Suominen P, Vehmaanperae J;
 XX
 DR WPI; 1996-141347/38.
 DR P-PSDB; AAY24480.
 XX
 XX Actinomadura xylanase sequences and method of use. - No abstract.
 PT
 XX
 PS Disclosure; Fig 13; 54pp; English.
 XX
 CC The present invention describes a culture medium obtained from the
 CC culture of a recombinant host cell that is not Actinomadura flexuosa and
 CC has been transformed with a vector encoding a protein having xylanase
 CC activity, where the protein comprises an amino acid sequence from a
 CC xylanolytic fragment of the amino acid sequence in AAY24480 or AAY24481.
 CC Also described is an enzyme preparation derived from the culture medium.
 CC The enzyme composition can be used in a method for biobleaching.
 CC modifying plant biomass properties, especially the reduction of lignin
 CC content in pulp and paper processing. The xylanases are hemicellulases
 CC which partially degrade the hemicellulose and enhance the extractability
 CC of lignins by conventional chemical bleaching of wood pulp. It can be
 CC used alone or as a supplement to other treatments that reduce lignin
 CC content of wood pulp, increase its drainability or decrease its water
 CC retention. The culture medium can be used directly without the need to
 CC purify the enzymes. Actinomadura flexuosa xylanases have a pH optimum
 CC and thermostability desirable for the biobleaching of wood pulp reducing
 CC the need to acidify the pulp prior to xylanase treatment. The xylanases
 CC partially degrade the hemicellulose in wood pulp which enhances the
 CC extractability of lignins by conventional bleaching chemicals and
 CC results in a lower consumption of bleaching chemicals reducing the
 CC formation of environmentally undesired organic compounds. The present
 CC sequence encodes Actinomadura sp. DSM43186 35 kDa xylanase.
 CC N.B. This sequence is indexed from US5935836 which is a treat as
 CC basic specification for F19503639.
 CC (Updated on 25-MAR-2003 to correct DR field.)
 XX
 XX Sequence 1375 BP; 265 A; 517 C; 400 G; 193 T; 0 other;
 SQ
 Query Match 18.0%; Score 119.4; DB 17; Length 1375;
 Best Local Similarity 57.1%; Pred. No. 2.1e-25;
 Matches 238; Conservative 0; Mismatches 176; Indels 3; Gaps 1;
 198 GTCCATAAACTACGGAGCGCAACTTCCAAACCAATGGTAATCGGTATTTATCGGTCTATGG 257

Db 608 GACCGTGACCTACAAACGCTCTCTTCAACCCGTCGGGTAAACGCTACCTCACGCTCTACGG 667
 Qy 258 TTGGACTGTTGACCCCTCTTGTGCAATATTATATTGTGACAGTTGGGGCAACTGGCGTCC 317
 Db 668 CTGACCCAGGAACCGGCTCGTCGAGTACTATCGTCGAGAGCTGGGGCACTTACCGGCC 727
 Qy 318 ACCAGGAGCAACGCTTAAGGGGACCATCCTCTTTGATGGAGGAACATATGATATCTACGA 377
 Db 728 CACCGG---CACCTACAGGGGACCGCTCACCGAGCGGGGACGTACGACATCTACGA 784
 Qy 378 GACTCTTAGAGTCAATCAACCTCTCATTAAGGGGATTCGCACATTTAAACAATATTGGAG 437
 Db 785 GACCTGGCGGTACAAACGCGCGCTCCATCGAGGGGACCCGCGACCTTCCAGCAGTTCTGGAG 844
 Qy 438 TGTTTCAAGATCGAAACGACGAGTGCGCAGCATTTCTGTGAGCAACCACTTTTAGAGCGTG 497
 Db 845 CGTCCGGCAGCAGAGCGGACCGGCGCACCATTCCCATCGGCAACCACTTCGACGCGCTG 904
 Qy 498 GGAATACTTAGGGATGAATATGGGAAATATGAAAGTCGCGCTTACTGTAGAGGCTA 557
 Db 905 GGGCGCGCGCATGAACCTGGGCGAGCCAGCACTACCGAGATCATGGGCGACCGAGGGCTA 964
 Qy 558 TCRAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGAATTAAACGGTAACCC 614
 Db 965 CCAGAGCAGCGGTAGTCTCACCGTCTCCATCAGCGGAGGGTGGCAACCCCGCGCAACCC 1021

RESULT 15
 AAT64930 standard; DNA; 1375 BP.
 ID AAT64930
 XX
 AC AAT64930;
 XX
 DT 25-MAR-2003 (updated)
 DT 02-APR-1998 (first entry)
 XX
 DE Actinomadura flexuosa 35 kDa (AM35) xylanase encoding DNA.
 XX
 KW Actinomadura flexuosa; xylanase; cellulase; recombinant; fungal host;
 KW pulp; paper industry; enzyme; bleaching; ss.
 XX
 OS Actinomadura flexuosa.
 FH Key Location/Qualifiers
 FT CDS 303..1337
 FT /*tag= a
 FT /product= "AM35 xylanase"
 FT mat_peptide 432..1334
 FT /*tag= b
 XX
 WO9727306-A1.
 XX
 PD 31-JUL-1997.
 XX
 XX 24-JAN-1997; 97WO-FI00037.
 XX
 XX 26-JAN-1996; 96US-0590563.
 XX
 PA (ALKO-) ALKO GROUP LTD.
 XX
 PI Maentylae A, Paloheimo M, Lantto R, Fagerstrom R, Lahtinen T;
 PI Suominen P, Vehmaanperae J;
 XX
 DR WPI; 1997-393693/36;
 DR P-PSDB; AAW23338.
 XX
 PT Production of bacterial proteins, especially xylanase(s) and
 PT cellulase(s) - by recombinant expression in a filamentous fungal
 PT host, useful particularly in the pulp and paper industries.
 XX
 PS Claim 7; Pages 65-67; 127pp; English.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: October 20, 2003, 10:40:37 ; Search time 2552 Seconds
(without alignments)
10628.167 Million cell updates/sec
Title: US-09-909-207-1
Perfect score: 663
Sequence: 1 CAATGTCACCGAATTC.....TAATTTGGATAAAACAAT 663

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 2888711 seqs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.vi.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.ly.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	663	100.0	663	6	A48222 Sequence 1
2	663	100.0	663	6	A48223 Sequence 2
3	663	100.0	663	6	AR193049 Sequence
4	663	100.0	663	6	AR193050 Sequence
5	663	100.0	744	6	A48225 Sequence 4
6	663	100.0	744	6	A48226 Sequence 5
7	663	100.0	744	6	AR193051 Sequence
8	663	100.0	744	6	AR193052 Sequence
9	663	100.0	1513	6	A48231 Sequence 10
10	663	100.0	1513	6	A48232 Sequence 11
11	663	100.0	1513	6	AR193055 Sequence
12	663	100.0	1513	6	AR193056 Sequence
13	634.2	95.7	744	6	A68006 Sequence 1
14	634.2	95.7	744	6	AR163110 Sequence
15	624.6	94.2	744	6	A45313 Sequence 18
16	624.6	94.2	744	6	AR117325 Sequence
17	507.8	76.6	744	6	A68016 Sequence 11
18	507.8	76.6	744	6	AR163117 Sequence
19	355.2	53.6	1531	1	BSXNY
20	355.2	53.6	2246	1	AB029319 Bacillus
21	280.4	42.3	1454	1	CLOXVNB
22	278.2	42.0	4128	1	AF047761 Clostridi
23	276.6	41.7	5774	1	AB010958 Clostridi
24	274.8	41.4	3493	1	CST508403 Clostridi
25	268.2	40.5	1070	1	BFXYNA
26	263.4	39.7	1011	1	AF490981 Bacillus
27	258.6	39.0	1789	1	AF326785 Bacillus
28	255.4	38.5	2364	1	CLOXYNA
29	255.4	38.5	2364	6	AR203507 Clostridium
30	255.2	38.5	1022	6	A42251 Clostridi
31	255.2	38.5	1022	6	A42285 Sequence 1
32	255.2	38.5	1022	6	AR127019 Sequence
33	255.2	38.5	1022	6	AR127049 Sequence
34	255.2	38.5	1022	6	AR220023 Sequence
35	255.2	38.5	1022	6	AR220053 Sequence
36	255.2	38.5	1022	6	AR221306 Sequence
37	255.2	38.5	1022	6	AR221336 Sequence
38	255	38.5	600	6	A42276 Sequence 26
39	255	38.5	600	6	A42284 Sequence 34
40	255	38.5	600	6	AR127041 Sequence
41	255	38.5	600	6	AR127048 Sequence
42	255	38.5	600	6	AR220045 Sequence
43	255	38.5	600	6	AR220052 Sequence
44	255	38.5	600	6	AR221328 Sequence
45	255	38.5	600	6	AR221335 Sequence

ALIGNMENTS

RESULT 1
A48222
LOCUS A48222
DEFINITION Sequence 1 from Patent (EP0698667).
ACCESSION A48222
VERSION A48222.1 GI:2302069
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 663)
AUTHORS De, B.E., Lahaye, A., Ledoux, P. and Detroz, R.
TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof
JOURNAL Patent: EP 0698667-A 1 28-FEB-1996;
linear PAT 07-MAR-1997

COMMENT SOLVAY (BE)
 Other publication BE 1008751 960702
 Other publication BE 1008570 960604
 Other publication BR 9503454 960305
 Other publication JP 8092284 960409
 Other publication FI 953578 960127
 Other publication CA 2154628 960127
 Other publication AU 2508695 960208.

FEATURES
 source
 1. .663
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 220 a 131 c 146 g 166 t

ORIGIN

Query Match 100.0%; Score 663; DB 6; Length 663;
 Best Local Similarity 100.0%; Pred. No. 1.5e-169;
 Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATCGTACCGACCAATTCATTTGGCAACACGATGCTATGATTGAAATTTGGAAA 60
 DB 1 CAATCGTACCGACCAATTCATTTGGCAACACGATGCTATGATTGAAATTTGGAAA 60
 QY 61 GATAGCGGTGCTGGGCAATGATTCATCATGCGGCTACCTTCAGTCCCAATGG 120
 DB 61 GATAGCGGTGCTGGGCAATGATTCATCATGCGGCTACCTTCAGTCCCAATGG 120
 QY 121 AACATGTTACCAACATATTTCCGTAAAGGTTAAATAATTCATAAACAACAACAC 180
 DB 121 AACATGTTACCAACATATTTCCGTAAAGGTTAAATAATTCATAAACAACAACAC 180
 QY 181 CAACAAGTTGTGTAACATGTCATTAACGAGACCACTTCACCAAGTGTATGCG 240
 DB 181 CAACAAGTTGTGTAACATGTCATTAACGAGACCACTTCACCAAGTGTATGCG 240
 QY 241 TATTTATGCTGTATGCTGAGCTGTGACCTCTTGTGTGAATATTAATTTGCCAC 300
 DB 241 TATTTATGCTGTATGCTGAGCTGTGACCTCTTGTGTGAATATTAATTTGCCAC 300
 QY 301 TGGGGCAACTGCGCTGCACACGAGACGACCTTAAGGGGACCATCTGTTATGAGGA 360
 DB 301 TGGGGCAACTGCGCTGCACACGAGACGACCTTAAGGGGACCATCTGTTATGAGGA 360
 QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTTCATTAAGGGGATGCGACA 420
 DB 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTTCATTAAGGGGATGCGACA 420
 QY 421 TTTAAACAATATTTGAGAGTGTGAAAGTGAAGACGAGAGTGGACGATTTCTGTCA 480
 DB 421 TTTAAACAATATTTGAGAGTGTGAAAGTGAAGACGAGAGTGGACGATTTCTGTCA 480
 QY 481 AACCACTTAGACGCTGGGAAAACTTAGGGATGAATATGAGGAAAAATGATGAAGTCGG 540
 DB 481 AACCACTTAGACGCTGGGAAAACTTAGGGATGAATATGAGGAAAAATGATGAAGTCGG 540
 QY 541 CTTACTGTGAAGGCTATCAAAAGTAGCGGAGAGTGTATATATGCAATACACTAAGA 600
 DB 541 CTTACTGTGAAGGCTATCAAAAGTAGCGGAGAGTGTATATATGCAATACACTAAGA 600
 QY 601 ATTAACGGTAACCCCTCTCACTAATATAGTATGAGAGATTAATTTGATTAATAAAC 660
 DB 601 ATTAACGGTAACCCCTCTCACTAATATAGTATGAGAGATTAATTTGATTAATAAAC 660
 QY 661 AAT 663
 DB 661 AAT 663

RESULT 2 A48223 663 bp DNA linear PART 07-MAR-1997
 LOCUS A48223
 DEFINITION Sequence 2 from Patent EP0698667.

ACCESSION A48223
 VERSION A48223.1 GI:2302070
 KEYWORDS
 SOURCE
 ORGANISM
 unclassified
 unclassified
 unclassified

REFERENCE
 1 (bases 1 to 663)
 De, B.E., Lahaye, A., Ledoux, P. and Detroz, R.
 Xylanase, microorganisms for its production, DNA molecules, process
 of preparation and use thereof
 Patent: EP 0698667-A 2 28-FEB-1996;

JOURNAL
 SOLVAY (BE)

COMMENT
 Other publication BE 1008751 960702
 Other publication BE 1008570 960604
 Other publication BR 9503454 960305
 Other publication JP 8092284 960409
 Other publication FI 953578 960127
 Other publication CA 2154628 960127
 Other publication AU 2508695 960208.

FEATURES
 source
 1. .663
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

CDS
 1. .663
 /note="unamed protein product; Protein sequence is in
 conflict with the conceptual translation"
 /codon_start=1
 /protein_id="CA03091.1"
 /db_xref="gi:2302071"
 /translation="QIYTDNSIGHNDYDFPMKDSGSGSTMILNHSPTSAOKNNN
 NILFRKGEKEIETHQOVGMSINYGAFQPNANALCYGVMTVDLVEYIDSWG
 NMRPGATPKGTITVDGTYDIEYELRNQPSIGIATFKQYVRSKRYSGISVS
 NHFRAMENTGLNMKMYEVALTVBEGYSSGANYSTLRINGNPSTISNDSITLD
 KNN"

mat_peptide 1. .663
 /product="unamed"

BASE COUNT 220 a 131 c 146 g 166 t

ORIGIN

Query Match 100.0%; Score 663; DB 6; Length 663;
 Best Local Similarity 100.0%; Pred. No. 1.5e-169;
 Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATCGTACCGACCAATTCATTTGGCAACACGATGCTATGATTGAAATTTGGAAA 60
 DB 1 CAATCGTACCGACCAATTCATTTGGCAACACGATGCTATGATTGAAATTTGGAAA 60
 QY 61 GATAGCGGTGCTGGGCAATGATTCATCATGCGGCTACCTTCAGTCCCAATGG 120
 DB 61 GATAGCGGTGCTGGGCAATGATTCATCATGCGGCTACCTTCAGTCCCAATGG 120
 QY 121 AACATGTTACCAACATATTTCCGTAAAGGTTAAATAATTCATAAACAACAACAC 180
 DB 121 AACATGTTACCAACATATTTCCGTAAAGGTTAAATAATTCATAAACAACAACAC 180
 QY 181 CAACAAGTTGTGTAACATGTCATTAACGAGACCACTTCACCAAGTGTATGCG 240
 DB 181 CAACAAGTTGTGTAACATGTCATTAACGAGACCACTTCACCAAGTGTATGCG 240
 QY 241 TATTTATGCTGTATGCTGAGCTGTGACCTCTTGTGTGAATATTAATTTGCCAC 300
 DB 241 TATTTATGCTGTATGCTGAGCTGTGACCTCTTGTGTGAATATTAATTTGCCAC 300
 QY 301 TGGGGCAACTGCGCTGCACACGAGACGACCTTAAGGGGACCATCTGTTATGAGGA 360
 DB 301 TGGGGCAACTGCGCTGCACACGAGACGACCTTAAGGGGACCATCTGTTATGAGGA 360
 QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTTCATTAAGGGGATGCGACA 420
 DB 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTTCATTAAGGGGATGCGACA 420
 QY 421 TTTAAACAATATTTGAGAGTGTGAAAGTGAAGACGAGAGTGGACGATTTCTGTCA 480

QY 421 TTTAAACATATGTGAGTGTTCGAAGATCGAAACGACGATGGCAAGATTCTGTACG 480
 DB 421 TTTAAACATATGTGAGTGTTCGAAGATCGAAACGACGATGGCAAGATTCTGTACG 480
 QY 461 AACCCCTTAGAGCGCTGGGAAAACCTTAGGATGATATCGGGGAAAATGATTAAGTCCGC 540
 DB 461 AACCCCTTAGAGCGCTGGGAAAACCTTAGGATGATATCGGGGAAAATGATTAAGTCCGC 540
 QY 541 CTTACTGTAGAGGCTATCAAGTAGCGGAGTGTCTAATGTATAGCAATACCTAAGA 600
 DB 541 CTTACTGTAGAGGCTATCAAGTAGCGGAGTGTCTAATGTATAGCAATACCTAAGA 600
 QY 601 ATTAACGGTAACCTCTCTCAACTATTAAGTATGACGAGACATACTTTGGATAAAC 660
 DB 601 ATTAACGGTAACCTCTCTCAACTATTAAGTATGACGAGACATACTTTGGATAAAC 660
 QY 661 AAT 663
 DB 661 AAT 663

RESULT 5

A48225 744 bp DNA linear PAT 07-MAR-1997
 LOCUS A48225
 DEFINITION Sequence 4 from Patent EP0698667.
 ACCESSION A48225
 VERSION A48225.1 GI:2302072
 KEYWORDS
 SOURCE

ORGANISM
 source

unidentified
 unidentified
 unclassified.

REFERENCE 1 (bases 1 to 744)
 AUTHORS De, B.E., Lahaye, A., Ledoux, P. and Detroz, R.
 TITLE Xylanase, microorganisms for its production, DNA molecules, process
 of preparation and use thereof
 JOURNAL Patent: EP 0698667-A 4 28-FEB-1996;
 SOLVAY (BE)

COMMENT Other publication BE 1008751 960702
 Other publication BE 1008570 960604
 Other publication BR 9503454 960305
 Other publication JP 8092284 960409
 Other publication FI 953578 960127
 Other publication CA 2154628 960127
 Other publication AU 2508695 960208.
 Location/Qualifiers

FEATURES
 source

1..744
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 246 a 144 c 160 g 194 t

Query Match 100.0%; Score 663; DB 6; Length 744;
 Beef Local Similarity 100.0%; Pred. No. 1.5e-169;
 Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAATCGTCACCGACATTCATTTGCAACACGATGGCTATGATTAATTTGGAAA 60
 DB 82 CAATCGTCACCGACATTCATTTGCAACACGATGGCTATGATTAATTTGGAAA 141
 QY 61 GATAGCGGTGGCTCTGGGCAATGATTTCTCATCATGCGGGTACCTTCCAGTCCCAATG 120
 DB 142 GATAGCGGTGGCTCTGGGCAATGATTTCTCATCATGCGGGTACCTTCCAGTCCCAATG 201
 QY 121 AACATGTTAAACAATATTTTCGTAAGGTAAGTAATTAATGAACACAAACACAC 180
 DB 202 AACATGTTAAACAATATTTTCGTAAGGTAAGTAATTAATGAACACAAACACAC 261
 QY 181 CAACAAGTTGTACATGTCATTAACCTACGAGCCCACTTCCAAACCAATGATATGCG 240
 DB 262 CAACAAGTTGTACATGTCATTAACCTACGAGCCCACTTCCAAACCAATGATATGCG 321

QY 241 TATTTAGCGTCTAATGTTGGACTGTGACCTCTTGTGCAATATATATTGTGACAGT 300
 DB 322 TATTTAGCGTCTAATGTTGGACTGTGACCTCTTGTGCAATATATATTGTGACAGT 381
 QY 301 TGGGGCAACCTGGCGTCCACGAGAGCAACGCTTAAGGGGACCATACACTGTTGATGAGGA 360
 DB 382 TGGGGCAACCTGGCGTCCACGAGAGCAACGCTTAAGGGGACCATACACTGTTGATGAGGA 441
 QY 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTTCATTAAGGGGATGGCCACA 420
 DB 442 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTTCATTAAGGGGATGGCCACA 501
 QY 421 TTTAAACAATTTGAGTGTTCGAAGATCGAAACGACAGATGGCAAGATTCTGTACG 480
 DB 502 TTTAAACAATTTGAGTGTTCGAAGATCGAAACGACAGATGGCAAGATTCTGTACG 561
 QY 481 AACCACTTAGAGCGTGGGAAAACCTTAGGATGATATCGGGGAAAATGATTAAGTCCGC 540
 DB 562 AACCACTTAGAGCGTGGGAAAACCTTAGGATGATATCGGGGAAAATGATTAAGTCCGC 621
 QY 541 CTTACTGTAGAGGCTATCAAGTAGCGGAGTGTCTAATGTATATAGCAATACCTAAGA 600
 DB 622 CTTACTGTAGAGGCTATCAAGTAGCGGAGTGTCTAATGTATATAGCAATACCTAAGA 681
 QY 601 ATTAACGGTAACCTCTCTCAACTATTAAGTATGACGAGACATACTTTGGATAAAC 660
 DB 682 ATTAACGGTAACCTCTCTCAACTATTAAGTATGACGAGACATACTTTGGATAAAC 741
 QY 661 AAT 663
 DB 742 AAT 744

RESULT 6

A48226 744 bp DNA linear PAT 07-MAR-1997
 LOCUS A48226
 DEFINITION Sequence 5 from Patent EP0698667.
 ACCESSION A48226
 VERSION A48226.1 GI:2302073
 KEYWORDS
 SOURCE

ORGANISM
 source

unidentified
 unidentified
 unclassified.

REFERENCE 1 (bases 1 to 744)
 AUTHORS De, B.E., Lahaye, A., Ledoux, P. and Detroz, R.
 TITLE Xylanase, microorganisms for its production, DNA molecules, process
 of preparation and use thereof
 JOURNAL Patent: EP 0698667-A 5 28-FEB-1996;
 SOLVAY (BE)

COMMENT Other publication BE 1008751 960702
 Other publication BE 1008570 960604
 Other publication BR 9503454 960305
 Other publication JP 8092284 960409
 Other publication FI 953578 960127
 Other publication CA 2154628 960127
 Other publication AU 2508695 960208.
 Location/Qualifiers

FEATURES
 source

1..744
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

1..>744
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CA03092.1"

CDS
 /translation="MRKKTLTLAFLVCFALTPARTIOAIVTDSISGHDGVDVE
 FMKSGSGGTITLHNGGTFSAQNNVNNILFRKSKREITTHQVGMISNYGANFO
 PNGNAVLCVGMIVDPLVEYIVDSWGNRPGATPKGTIVDGTVDIYETLRVNP
 SIKIATFKQVSWRSKRTSGTISVSWHFRPAMENLGMNMKMEVALTVBGYOSGS
 ANYVSNTRLRINGNPLSTISNDSITLIDKN"
 sig_peptide 1..81
 mat_peptide 82..744

BASE COUNT		246 a	144 c	160 g	194 t
ORIGIN		/product="unnamed"			
Query Match		100.0%; Score 663; DB 6; Length 744;			
Best Local Similarity		100.0%; Pred. No. 1.5e-169;			
Matches .663; Conservative		0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CAAAATCGTCACCGACAATTC	CAATTCGCGCAACACACGATG	CGCTATGATTAATGAATTTGGAAA	60
Db	82	CAAAATCGTCACCGACAATTC	CAATTCGCGCAACACACGATG	CGCTATGATTAATGAATTTGGAAA	141
Qy	61	GATAGCGTGGCTCTGGGACAAT	GAATTCCTCAATCATGCGCGTAC	CGTTCAGTGCCTCAATGG	120
Db	142	GATAGCGTGGCTCTGGGACAAT	GAATTCCTCAATCATGCGCGTAC	CGTTCAGTGCCTCAATGG	201
Qy	121	AACAATGTTAAACAACATATT	TCCTCGTAAAGGTAAATAATTC	CAATGAACAACAACACAC	180
Db	202	AACAATGTTAAACAACATATT	TCCTCGTAAAGGTAAATAATTC	CAATGAACAACAACACAC	261
Qy	181	CAACAAGTTGGTAACATGTCC	ATAAATCTACGAGGCAACTTC	CAACCAAAATGGTAATGCG	240
Db	262	CAACAAGTTGGTAACATGTCC	ATAAATCTACGAGGCAACTTC	CAACCAAAATGGTAATGCG	321
Qy	241	TATTTATCGCTCTATGGTTGG	ACTGTTGACCCCTCTTGTGCA	ATATTATTTGTCGACAGT	300
Db	322	TATTTATCGCTCTATGGTTGG	ACTGTTGACCCCTCTTGTGCA	ATATTATTTGTCGACAGT	381
Qy	301	TGGGGCAACTGGCGTCCACG	AGGACAAACGCTTAAGGGGAC	CACTCTGTTGATGGAGGA	360
Db	382	TGGGGCAACTGGCGTCCACG	AGGACAAACGCTTAAGGGGAC	CACTCTGTTGATGGAGGA	441
Qy	361	ACATATGATATCTACGAGACT	CTTAGAGTCAATCAACCTCTC	ATTAAAGGGGATTCGCCACA	420
Db	442	ACATATGATATCTACGAGACT	CTTAGAGTCAATCAACCTCTC	ATTAAAGGGGATTCGCCACA	501
Qy	421	TTTAAACAATATTTGGAGTGT	TCGAAGATCGAAACGACAGT	GGCAGCATCTTCTGTGACG	480
Db	502	TTTAAACAATATTTGGAGTGT	TCGAAGATCGAAACGACAGT	GGCAGCATCTTCTGTGACG	561
Qy	481	AACCACTTTAGAGCGTGGGAAA	CTTAGGGATGAATATGGGGAAA	TAATGTATGAAGTCGCG	540
Db	562	AACCACTTTAGAGCGTGGGAAA	CTTAGGGATGAATATGGGGAAA	TAATGTATGAAGTCGCG	621
Qy	541	CTTACTGTAGAGGCTATCAAG	TAGCGAGTCTTAATGTATAG	CAATACACTAAGA	600
Db	622	CTTACTGTAGAGGCTATCAAG	TAGCGAGTCTTAATGTATAG	CAATACACTAAGA	681
Qy	601	ATTAAACGGTAACCTCTCTCA	ACTTATAGTAAATGACGAGAC	ATAACTTTTGGATAAAAAAC	660
Db	682	ATTAAACGGTAACCTCTCTCA	ACTTATAGTAAATGACGAGAC	ATAACTTTTGGATAAAAAAC	741
Qy	661	AAT 663			
Db	742	AAT 744			

RESULT 7	AR193051	Sequence 4 from patent US 6346407.	744 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR193051	Sequence 4 from patent US 6346407.	744 bp	DNA	linear	PAT 20-APR-2002
DEFINITION	Sequence 4 from patent US 6346407.	744 bp	DNA	linear	PAT 20-APR-2002	
ACCESSION	AR193051	Sequence 4 from patent US 6346407.	744 bp	DNA	linear	PAT 20-APR-2002
VERSION	AR193051.1	GI:20239016				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 744)					
AUTHORS	De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.					
TITLE	Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter					
JOURNAL	Patent: US 6346407-A 4 12-FEB-2002;					
FEATURES	Location/Qualifiers					

RESULT 8	AR193052	Sequence 5 from patent US 6346407.	744 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR193052	Sequence 5 from patent US 6346407.	744 bp	DNA	linear	PAT 20-APR-2002
DEFINITION	Sequence 5 from patent US 6346407.	744 bp	DNA	linear	PAT 20-APR-2002	
ACCESSION	AR193052	Sequence 5 from patent US 6346407.	744 bp	DNA	linear	PAT 20-APR-2002
VERSION	AR193052.1	GI:20239017				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 744)					
AUTHORS	De Buyl, E., Lahaye, A., Ledoux, P. and Detroz, R.					
TITLE	Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter					
JOURNAL	Patent: US 6346407-A 5 12-FEB-2002;					
FEATURES	Location/Qualifiers					

FEATURES
source
Location/Qualifiers
1. 744
/organism="unknown"
BASE COUNT 246 a 144 c 160 g 194 t
ORIGIN

Query Match 100.0%; Score 663; DB 6; Length 744;
Best Local Similarity 100.0%; Pred. No. 1.5e-169;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAATCGTCACCGCAATTCCTGATGCAACACGATGGCTATGATTAATTTGGAAA 60
2 CAATCGTCACCGCAATTCCTGATGCAACACGATGGCTATGATTAATTTGGAAA 141
3 GATACGGTGGCTCTGGGCAATGATTCATCAATGGCGATGCTGATGCTCCAAATGG 201
4 GATACGGTGGCTCTGGGCAATGATTCATCAATGGCGATGCTGATGCTCCAAATGG 201
5 AACATGTTAAACAATATATATTCGTTAAAGGTTAAATTCATGAAACAAACACAC 180
6 AACATGTTAAACAATATATATTCGTTAAAGGTTAAATTCATGAAACAAACACAC 261
7 CAACAAGTTGGTAAATGTCATTAATCTAGGAGCCAACTTCCAAACCAATGGTATGG 240
8 CAACAAGTTGGTAAATGTCATTAATCTAGGAGCCAACTTCCAAACCAATGGTATGG 321
9 TATTTATGGCTGATGTTGATGATGTCGATGCTTGTGGAATATATATGTCACAGT 300
10 TATTTATGGCTGATGTTGATGATGTCGATGCTTGTGGAATATATATGTCACAGT 381
11 TGGGGCAATGGCGTCCACAGAGACCAAGCTTAAGGGGACATCACTGTTGATGAGGA 360
12 TGGGGCAATGGCGTCCACAGAGACCAAGCTTAAGGGGACATCACTGTTGATGAGGA 441
13 TGGGGCAATGGCGTCCACAGAGACCAAGCTTAAGGGGACATCACTGTTGATGAGGA 441
14 ACATATGATATCTAGAGACTCTTAGAGTCATCAACCTCCATTAAAGGGATGGCACA 420
15 ACATATGATATCTAGAGACTCTTAGAGTCATCAACCTCCATTAAAGGGATGGCACA 501
16 TTTAAACAATATGAGTGTTCGAAAGTCGAAACGACAGTGGGACGATTTCTGTCAAC 480
17 TTTAAACAATATGAGTGTTCGAAAGTCGAAACGACAGTGGGACGATTTCTGTCAAC 561
18 AACCTTTAGAGCGTGGGAAAACCTTAGGAGTGAATATGAGGAAAATGATAGTCGG 540
19 AACCTTTAGAGCGTGGGAAAACCTTAGGAGTGAATATGAGGAAAATGATAGTCGG 621
20 CTTACTGTAGAGGCTATCAAGTAGACGGAAGTGTATATAGCAATACATAGA 600
21 CTTACTGTAGAGGCTATCAAGTAGACGGAAGTGTATATAGCAATACATAGA 681
22 ATTAAACGTTAACCTCTCTCAACTATAGTAATGACGAGACATTAATTTGATAAAAAC 660
23 ATTAAACGTTAACCTCTCTCAACTATAGTAATGACGAGACATTAATTTGATAAAAAC 741
24 AAT 663
25 AAT 744

RESULT 9
LOCUS A48231 1513 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 10 from Patent EP0698667.
ACCESSION A48231
VERSION A48231.1 GI:2302078
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1513)
AUTHORS De, B.E., Lahaye, A., Ledoux, P. and Detroz, R.
TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof

JOURNAL Patent: EP 0698667-A 10 28-FEB-1996;
COMMENT SOLWAY (BE)
Other publication BE 1008751 960702
Other publication BE 1008570 960604
Other publication BR 9503454 960305
Other publication JP 8092284 960409
Other publication FI 953578 960127
Other publication CA 2154628 960127
Other publication AU 2508695 960208.

FEATURES
source
Location/Qualifiers
1. 1513
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 500 a 282 c 271 g 460 t
ORIGIN

Query Match 100.0%; Score 663; DB 6; Length 1513;
Best Local Similarity 100.0%; Pred. No. 1.4e-169;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAATCGTCACCGCAATTCCTGATGCAACACGATGGCTATGATTAATTTGGAAA 60
2 CAATCGTCACCGCAATTCCTGATGCAACACGATGGCTATGATTAATTTGGAAA 760
3 GATACGGTGGCTCTGGGCAATGATTCATCAATGGCGATGCTGATGCTCCAAATGG 120
4 GATACGGTGGCTCTGGGCAATGATTCATCAATGGCGATGCTGATGCTCCAAATGG 820
5 AACATGTTAAACAATATATATTCGTTAAAGGTTAAATTCATGAAACAAACACAC 180
6 AACATGTTAAACAATATATATTCGTTAAAGGTTAAATTCATGAAACAAACACAC 880
7 CAACAAGTTGGTAAATGTCATTAATCTAGGAGCCAACTTCCAAACCAATGGTATGG 240
8 CAACAAGTTGGTAAATGTCATTAATCTAGGAGCCAACTTCCAAACCAATGGTATGG 940
9 TATTTATGGCTGATGTTGATGATGTCGATGCTTGTGGAATATATATGTCACAGT 300
10 TATTTATGGCTGATGTTGATGATGTCGATGCTTGTGGAATATATATGTCACAGT 1000
11 TGGGGCAATGGCGTCCACAGAGACCAAGCTTAAGGGGACATCACTGTTGATGAGGA 360
12 TGGGGCAATGGCGTCCACAGAGACCAAGCTTAAGGGGACATCACTGTTGATGAGGA 1060
13 ACATATGATATCTAGAGACTCTTAGAGTCATCAACCTCCATTAAAGGGATGGCACA 420
14 ACATATGATATCTAGAGACTCTTAGAGTCATCAACCTCCATTAAAGGGATGGCACA 1120
15 TTTAAACAATATGAGTGTTCGAAAGTCGAAACGACAGTGGGACGATTTCTGTCAAC 480
16 TTTAAACAATATGAGTGTTCGAAAGTCGAAACGACAGTGGGACGATTTCTGTCAAC 1180
17 AACCTTTAGAGCGTGGGAAAACCTTAGGAGTGAATATGAGGAAAATGATAGTCGG 540
18 AACCTTTAGAGCGTGGGAAAACCTTAGGAGTGAATATGAGGAAAATGATAGTCGG 1240
19 CTTACTGTAGAGGCTATCAAGTAGACGGAAGTGTATATAGCAATACATAGA 600
20 CTTACTGTAGAGGCTATCAAGTAGACGGAAGTGTATATAGCAATACATAGA 1300
21 ATTAAACGTTAACCTCTCTCAACTATAGTAATGACGAGACATTAATTTGATAAAAAC 660
22 ATTAAACGTTAACCTCTCTCAACTATAGTAATGACGAGACATTAATTTGATAAAAAC 1360
23 AAT 663
24 AAT 1363

RESULT 10
LOCUS A48232 1513 bp DNA linear PAT 07-MAR-1997
A48232

DEFINITION Sequence 11 from Patent EP0698667.
 ACCESSION A48232
 VERSION A48232.1 GI:2302079
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (bases 1 to 1513)
 AUTHORS De,B.E., Lahaye,A., Ledoux,P. and Detroz,R.
 TITLE Xylanase, microorganisms for its production, DNA molecules, process of preparation and use thereof
 JOURNAL Patent: EP 0698667-A 11 28-FEB-1996;
 SOLVAY (BE)
 COMMENT Other publication BE 1008751 960702
 Other publication BE 1008570 960604
 Other publication BR 9503454 960305
 Other publication JP 8092284 960409
 Other publication FI 953578 960127
 Other publication CA 2154628 960127
 Other publication AU 2508695 960208.
 FEATURES
 source
 1. .1513
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 CDS
 620. .1366
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAA03094.1"
 /db_xref="GI:2302080"
 /translation="MRQKKLLIAPLVCFALTPAEIIQAIQIVTDSIGNHDGYDVE
 FWKDSGGTLMHKGKFFSAQNNVNNILPRKKFNETQHQVGNMSINYANFQ
 PNGAYLCVYGTVDPLVEYYIVDSWGNWRPPGATPKGTITVDGGTIDYVETLRVNP
 SKIGAIYFKQWVSRRSRKTSGTISVSNHFRAWENLGNWGMKMEVALTVBEGYQSSGS
 ANVYSNTRLRINGNPLSTISNDSITLDKN"
 sig_peptide
 620. .700
 mat_peptide
 701. .1363
 /product="unnamed"
 BASE COUNT 500 a 282 c 271 g 460 t
 ORIGIN
 Query Match 100.0%; Score 663; DB 6; Length 1513;
 Best Local Similarity 100.0%; Pred. No. 1.4e-169;
 Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CAATCGTCACCGACAATTCATGGCAACACGATGGCTATGATTATGAATTTTGGAAA 60
 Db 701 CAATCGTCACCGACAATTCATGGCAACACGATGGCTATGATTATGAATTTTGGAAA 760
 Qy 61 GATAGCGGTGGCTCTGGGACAATGATTCCTCAATCATGCGCGTACGTTTCAGTGCCTCAATGG 120
 Db 761 GATAGCGGTGGCTCTGGGACAATGATTCCTCAATCATGCGCGTACGTTTCAGTGCCTCAATGG 820
 Qy 121 AACAAATGTTAACACATATTTATTCGCTAAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 180
 Db 821 AACAAATGTTAACACATATTTATTCGCTAAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 880
 Qy 181 CAACAAGTTGGTAACATGTCCTAAACCTACGAGGACCACTTCCAAACCAATGGTAATGCG 240
 Db 881 CAACAAGTTGGTAACATGTCCTAAACCTACGAGGACCACTTCCAAACCAATGGTAATGCG 940
 Qy 241 TATTTATGCGTCTATGGTTGGACTGTTGACCTCTTGTGCAATATTTATTTGTCGACAGT 300
 Db 941 TATTTATGCGTCTATGGTTGGACTGTTGACCTCTTGTGCAATATTTATTTGTCGACAGT 1000
 Qy 301 TGGGGCAACTGGCTCCACGAGGACCACTTAAGGGACCACTGTTGATGGAGGA 360
 Db 1001 TGGGGCAACTGGCTCCACGAGGACCACTTAAGGGACCACTGTTGATGGAGGA 1060
 Qy 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTTGCCACA 420
 Db 1061 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTTGCCACA 1120

Qy. 421 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCACGATTTCTGTCTAGC 480
 Db 1121 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCACGATTTCTGTCTAGC 1180
 Qy 481 AACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATATGGGAAAATCTATGAAATCTGAGTCCGC 540
 Db 1181 AACCACTTTAGAGCGTGGGAAAACCTTAGGGATGAATATGGGAAAATCTATGAAATCTGAGTCCGC 1240
 Qy 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTTAATGTATATAGCAATACACTAAGA 600
 Db 1241 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTTAATGTATATAGCAATACACTAAGA 1300
 Qy 601 ATTAACGGTAACCTCTCTCAACTATTAGTAATACGAGAGCATAACTTTTGGATAAAAAAC 660
 Db 1301 ATTAACGGTAACCTCTCTCAACTATTAGTAATACGAGAGCATAACTTTTGGATAAAAAAC 1360
 Qy 661 AAT 663
 Db 1361 AAT 1363
 RESULT 11
 ARI93055
 LOCUS ARI93055 1513 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 10 from patent US 6346407.
 ACCESSION ARI93055
 VERSION ARI93055.1 GI:20239020
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1513)
 AUTHORS De Buyl,E., Lahaye,A., Ledoux,P. and Detroz,R.
 TITLE Xylanase, microorganisms producing it, DNA molecules, methods for preparing this xylanase and uses of the latter
 JOURNAL Patent: US 6346407-A 10 12-FEB-2002;
 FEATURES
 source
 1. .1513
 /organism="unknown"
 BASE COUNT 500 a 282 c 271 g 460 t
 ORIGIN
 Query Match 100.0%; Score 663; DB 6; Length 1513;
 Best Local Similarity 100.0%; Pred. No. 1.4e-169;
 Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CAATCGTCACCGACAATTCATGGCAACACGATGGCTATGATTATGAATTTTGGAAA 60
 Db 701 CAATCGTCACCGACAATTCATGGCAACACGATGGCTATGATTATGAATTTTGGAAA 760
 Qy 61 GATAGCGGTGGCTCTGGGACAATGATTCCTCAATCATGCGCGTACGTTTCAGTGCCTCAATGG 120
 Db 761 GATAGCGGTGGCTCTGGGACAATGATTCCTCAATCATGCGCGTACGTTTCAGTGCCTCAATGG 820
 Qy 121 AACAAATGTTAACACATATTTATTCGCTAAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 180
 Db 821 AACAAATGTTAACACATATTTATTCGCTAAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGG 880
 Qy 181 CAACAAGTTGGTAACATGTCCTAAACCTACGAGGACCACTTCCAAACCAATGGTAATGCG 240
 Db 881 CAACAAGTTGGTAACATGTCCTAAACCTACGAGGACCACTTCCAAACCAATGGTAATGCG 940
 Qy 241 TATTTATGCGTCTATGGTTGGACTGTTGACCTCTTGTGCAATATTTATTTGTCGACAGT 300
 Db 941 TATTTATGCGTCTATGGTTGGACTGTTGACCTCTTGTGCAATATTTATTTGTCGACAGT 1000
 Qy 301 TGGGGCAACTGGCTCCACGAGGACCACTTAAGGGACCACTGTTGATGGAGGA 360
 Db 1001 TGGGGCAACTGGCTCCACGAGGACCACTTAAGGGACCACTGTTGATGGAGGA 1060
 Qy 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTTGCCACA 420
 Db 1061 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGGATTTGCCACA 1120

```

QY 421 TTTAAACAATATTTGAGTGTTCGAAGATGAAACGACGAGTGGACGATTTCTGTACG 480
DB 1121 TTTAAACAATATTTGAGTGTTCGAAGATGAAACGACGAGTGGACGATTTCTGTACG 1180
QY 481 AACCACTTTAGAGCGTGGGAAAATTAGGAGTGAATATGGGAAAAATGATAGAGTCGCG 540
DB 1181 AACCACTTTAGAGCGTGGGAAAATTAGGAGTGAATATGGGAAAAATGATAGAGTCGCG 1240
QY 541 CTTACTGTAGAGCGTATCAAGTAGCGGAAGTGTCTATGTATATAGCAATACACTAAGA 600
DB 1241 CTTACTGTAGAGCGTATCAAGTAGCGGAAGTGTCTATGTATATAGCAATACACTAAGA 1300
QY 601 ATTTAAGGTAAACCTCTCTCAACTATTTAGTAATGACGAGACATTAATTGGATAAAAC 660
DB 1301 ATTTAAGGTAAACCTCTCTCAACTATTTAGTAATGACGAGACATTAATTGGATAAAAC 1360
QY 661 AAT 663
DB 1361 AAT 1363

RESULT 12
A6193056 1513 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 11 from patent US 6346407.
DEFINITION A6193056
ACCESSION A6193056
VERSION A6193056.1 GI:20239021
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1513)
AUTHORS De Buy, E., Lahaye, A., Ledoux, P. and Detroz, R.
TITLE Xylanase, microorganisms producing it, DNA molecules, methods for
JOURNAL preparing this xylanase and uses of the latter
FEATURES
    source          location/Qualifiers
BASE COUNT 500 a 282 c 271 g 460 t
ORIGIN
Query Match 100.0%; Score 663; DB 6; Length 1513;
Best Local Similarity 100.0%; Pred. No. 1.4e-169;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATCGTACCGCAATTCATTTGGCAACGAGATGGCTATGATGAAATTTGGAAA 60
DB 701 CAAATCGTACCGCAATTCATTTGGCAACGAGATGGCTATGATGAAATTTGGAAA 760
QY 61 GATAGCGGTGCTCTGGGCAATGATTTCAATCATGAGCGGTACGTTCCAGTCCCAATG 120
DB 761 GATAGCGGTGCTCTGGGCAATGATTTCAATCATGAGCGGTACGTTCCAGTCCCAATG 820
QY 121 AACATGTTAACACATATTTATTCGTAAGGTAATAAATTCATGAAACACAAACAC 180
DB 821 AACATGTTAACACATATTTATTCGTAAGGTAATAAATTCATGAAACACAAACAC 880
QY 181 CAAAGAGTGTGAACATGTCATTAACCTACGAGCCCACTTCCAAACCAATGGTATGG 240
DB 881 CAAAGAGTGTGAACATGTCATTAACCTACGAGCCCACTTCCAAACCAATGGTATGG 940
QY 241 TATTTATGCGTCTATGAGTGTGACCTCTCTTGTGAATATATATTTGTGCAAGT 300
DB 941 TATTTATGCGTCTATGAGTGTGACCTCTCTTGTGAATATATATTTGTGCAAGT 1000
QY 301 TGGGGCACTGGCGTCCACGAGAGCAACGCTTAAGGGGACATCACTGTTGATGAGCA 360
DB 1001 TGGGGCACTGGCGTCCACGAGAGCAACGCTTAAGGGGACATCACTGTTGATGAGCA 1060
QY 361 ACATATGATATCTACGAGACTCTTAAGTCAATCAACCTCCATTAAGGGGATGGCACA 420

```

```

DB 1061 ACATATGATATCTACGAGACTCTTAAGTCAATCAACCTCCATTAAGGGGATGGCACA 1120
QY 421 TTTAAACAATATTTGAGTGTTCGAAGATGAAACGACGAGTGGACGATTTCTGTACG 480
DB 1121 TTTAAACAATATTTGAGTGTTCGAAGATGAAACGACGAGTGGACGATTTCTGTACG 1180
QY 481 AACCACTTTAGAGCGTGGGAAAATTAGGAGTGAATATGGGAAAAATGATAGAGTCGCG 540
DB 1181 AACCACTTTAGAGCGTGGGAAAATTAGGAGTGAATATGGGAAAAATGATAGAGTCGCG 1240
QY 541 CTTACTGTAGAGCGTATCAAGTAGCGGAAGTGTCTATGTATATAGCAATACACTAAGA 600
DB 1241 CTTACTGTAGAGCGTATCAAGTAGCGGAAGTGTCTATGTATATAGCAATACACTAAGA 1300
QY 601 ATTTAAGGTAAACCTCTCTCAACTATTTAGTAATGACGAGACATTAATTGGATAAAAC 660
DB 1301 ATTTAAGGTAAACCTCTCTCAACTATTTAGTAATGACGAGACATTAATTGGATAAAAC 1360
QY 661 AAT 663
DB 1361 AAT 1363

```

```

RESULT 13
A68006 744 bp DNA linear PAT 05-MAY-1999
LOCUS Sequence 1 from Patent WO9743409.
DEFINITION A68006
ACCESSION A68006
VERSION A68006.1 GI:4756810
KEYWORDS
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 744)
AUTHORS Dalboege, H., Diderichsen, B., Sandal, T. and Kauppinen, S.
TITLE METHOD OF PROVIDING NOVEL DNA SEQUENCES
JOURNAL Patent: WO 9743409-A 1 20-NOV-1997;
NOVONORDISK AS (DK)
FEATURES
    source          location/Qualifiers
BASE COUNT 249 a 146 c 157 g 192 t
ORIGIN
Query Match 95.7%; Score 634.2; DB 6; Length 744;
Best Local Similarity 97.3%; Pred. No. 1e-161;
Matches 645; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

CDS
1..5744
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42305.1"
/db_xref="GI:4756810"
/translation="MRKKLTFILAFVCPALTPAEIIQAOIVTDSIGNHGYDYE
FWKDSGSGTMI LNHGTFSAQMNNVNII LPRKGRKREYTHQOVGNMSINYGANFQ
ENKSNVILCVYGTVDPLVEYIYVDSGMMWRPGATPKITIVDGGTVDIYETLRVNP
SIKGIATKQYWSVRSKRTSGTISVSHRFBEMILGNMGMKRYEVALTVBEGYQSSGS
ANVSNLTRINGNPLSTISNDKSTILDGN"
Location/Qualifiers
1..744
/organism="unidentified"
/mol_type="genomic DNA"
/strain="BACILLUS SP. AC13, NCIMB NO. 40482"
/db_xref="E:taxon:32644"
1..5744
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB42305.1"
/db_xref="GI:4756810"
/translation="MRKKLTFILAFVCPALTPAEIIQAOIVTDSIGNHGYDYE
FWKDSGSGTMI LNHGTFSAQMNNVNII LPRKGRKREYTHQOVGNMSINYGANFQ
ENKSNVILCVYGTVDPLVEYIYVDSGMMWRPGATPKITIVDGGTVDIYETLRVNP
SIKGIATKQYWSVRSKRTSGTISVSHRFBEMILGNMGMKRYEVALTVBEGYQSSGS
ANVSNLTRINGNPLSTISNDKSTILDGN"

```



```

/translation="MSOKRLTILNPSLPALTLPARISOALVTMDNSIATRGVDEE
WDSGSGTMIINHSGTSAQNNNNNTLPRKGGKFNETQTHQVGNMSINGANFOP
NGNAYLCVYGMVDPLVEYIYVDSWGNMRPGATPKGTTVDGTYDIYETLRVNOP
IKGIATFKQYMSVRSKRISGTISVSNHFRAMENLGMNMGRKYEVALTVEGYOSSGA
NYSNTRLRINGNPLSTISNNESITLTKNN"
BASE COUNT      248 a      145 c      159 g      192 t
ORIGIN

```

```

Query Match      94.2%; Score 624.6; DB 6; Length 744;
Best Local Similarity 96.4%; Pred.No. 4.2e-159;
Matches 639; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```

QY	1	CAAAATCGTCACCGAGCAATTCCTTGGCAACCAAGTGTATGATTAATTTGGAAA	60
DB	79	CAAAATCGTCACCGAGCAATTCCTTGGCAACCAAGTGTATGATTAATTTGGAAA	138
QY	61	GATACGGGTGGCTTGGGCAATGATTCATCATGCGGTACGTTCAGTCCCAATGG	120
DB	139	GATACGGGTGGCTTGGGCAATGATTCATCATGCGGTACGTTCAGTCCCAATGG	198
QY	121	AACAAATGTTAACACATATTTATTCGTAAAGGTAAAAATTCATGAAACACAAACAC	180
DB	199	AATAATGTTAACATATATTTATTCGTAAAGGTAAAAATTCATGAAACACAAACAC	258
QY	181	CAACAAATGTTAACATATTCCTTAACTACGAGCCCACTTCCAAACCAATGTATGG	240
DB	259	CAACAAATGTTAACATATTCCTTAACTACGAGCCCACTTCCAAACCAATGTATGG	318
QY	241	TATTTATGCGGTATGATGTTGACCTGTTGACCTTGTGCAATATTTATGTCACAGT	300
DB	319	TATTTATGCGGTATGATGTTGACCTGTTGACCTTGTGCAATATTTATGTCACAGT	378
QY	301	TGGGGCAACTGGCGTCCACAGAGCAACGCTTAAGGGGACCATCATCTGTTGATGAGGA	360
DB	379	TGGGGCAACTGGCGTCCACAGAGCAACGCTTAAGGGGACCATCATCTGTTGATGAGGA	438
QY	361	ACATATGATATTTAGAGAGCTTTAGAGCATCAACCTCCATTAAGGGGATGGCCACA	420
DB	439	ACATATGATATTTAGAGAGCTTTAGAGCATCAACCTCCATTAAGGGGATGGCCACA	498
QY	421	TTTAAACAATATTTAGAGAGTTCGAAGATGAAACGACAGATGGACAGATTTCTGTACAGC	480
DB	499	TTTAAACAATATTTAGAGAGTTCGAAGATGAAACGACAGATGGACAGATTTCTGTACAGC	558
QY	481	AACCACTTTAGAGCGTGGGAAAACTTAGGGATGAATATGGGAAAAATGTATGAAGTCGCG	540
DB	559	AACCACTTTAGAGCGTGGGAAAACTTAGGGATGAATATGGGAAAAATGTATGAAGTCGCG	618
QY	541	CTTACTGTAGAGCGTATCAAGTAGCGGAGTGCCTAATGTATATGCAATACACTAAGA	600
DB	619	CTTACTGTAGAGCGTATCAAGTAGCGGAGTGCCTAATGTATATGCAATACACTAAGA	678
QY	601	ATTAAAGGTAAACCTCTCTCAACTATTAAGTAAGACGAGCATAACTTTGGATAAAAAC	660
DB	679	ATTAAAGGTAAACCTCTCTCAACTATTAAGTAAGACGAGCATAACTTTGGATAAAAAC	738
QY	661	AAT 663	
DB	739	AAT 741	

Search completed: October 20, 2003, 11:23:26
Job time : 2555 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 20, 2003, 11:28:06 ; Search time 1595 Seconds
(without alignments)
1091.613 Million cell updates/sec

Title: US-09-909-207-1

Perfect score: 663

Sequence: 1 CAATCGTCACGCAATTC.....TAATTCGATAAACAAT 663

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	663	100.0	663	10	US-09-909-207-1
2	663	100.0	663	10	US-09-909-207-2
3	663	100.0	744	10	US-09-909-207-4
4	663	100.0	744	10	US-09-909-207-5
5	663	100.0	1513	10	US-09-909-207-10
6	663	100.0	1513	10	US-09-909-207-11
7	119.4	18.0	1375	9	US-09-770-621-1
8	119.4	18.0	1375	12	US-10-286-993-1
9	115.2	17.4	596	12	US-10-307-441-39
10	111.4	16.8	942	14	US-10-213-990-71
11	110	16.6	1002	14	US-10-213-990-70
12	108.8	16.4	705	14	US-10-213-990-68
13	106	16.0	2898	14	US-10-299-393-1
14	104.8	15.8	712	14	US-10-213-990-64
15	103.6	15.6	666	14	US-10-213-990-65
16	96.8	14.6	739	14	US-10-213-990-67

17	86.8	13.1	678	11	US-09-803-454-3	Sequence 3, Appli
18	83.2	12.5	645	12	US-10-237-386-10	Sequence 10, Appl
19	83.2	12.5	657	12	US-10-237-386-11	Sequence 11, Appl
20	79	11.9	2225	11	US-09-790-070A-8	Sequence 8, Appli
21	72.4	10.9	588	12	US-10-237-386-9	Sequence 9, Appli
22	72.4	10.9	983	10	US-09-467-368-1	Sequence 1, Appli
23	54.2	8.2	2054	12	US-10-419-969-5	Sequence 5, Appli
24	50.4	7.6	643	12	US-10-340-860A-39	Sequence 39, Appl
25	41	6.2	74	12	US-10-307-441-24	Sequence 24, Appl
26	40	6.0	557	10	US-09-970-616-1	Sequence 1, Appli
27	38	5.7	6123	12	US-10-311-455-793	Sequence 793, App
28	37.6	5.7	10528	12	US-10-311-455-308	Sequence 308, App
29	36	5.4	787	13	US-10-027-632-13285	Sequence 13285, A
30	36	5.4	8622	12	US-10-311-455-2116	Sequence 2116, Ap
31	35.8	5.4	8880	12	US-09-840-743-10	Sequence 10, Appl
32	35.8	5.4	3673778	12	US-10-312-841-2	Sequence 2, Appli
33	35.6	5.4	478	13	US-10-027-632-63151	Sequence 63151, A
34	35.6	5.4	478	13	US-10-027-632-297516	Sequence 297516,
35	35.6	5.4	5333	8	US-08-781-986A-170	Sequence 170, App
36	35.2	5.3	623	13	US-10-027-632-212128	Sequence 212128,
37	35.2	5.3	980	13	US-10-027-632-9387	Sequence 9387, Ap
38	35.2	5.3	107820	12	US-09-792-616-1	Sequence 1, Appli
39	35	5.3	611	13	US-10-027-632-199892	Sequence 199892,
40	35	5.3	611	13	US-10-027-632-199893	Sequence 199893,
41	35	5.3	10832	12	US-10-311-455-1872	Sequence 1872, Ap
42	34.8	5.2	14307	12	US-10-311-455-701	Sequence 701, App
43	34.8	5.2	17527	12	US-10-311-455-1405	Sequence 1405, Ap
44	34.8	5.2	3673778	12	US-10-312-841-1	Sequence 1, Appli
45	34.6	5.2	5914	12	US-10-311-455-2141	Sequence 2141, Ap

ALIGNMENTS

RESULT 1

US-09-909-207-1
; Sequence 1, Application US/09909207
; Patent No. US20020115181A1
; GENERAL INFORMATION:

APPLICANT: ANDRE LAHAYE
ERIC DE BUYL
PIERRE LEDOUX
RENE DETROZ

TITLE OF INVENTION: Xylanase, microorganisms produced it,
DNA molecule, processes for preparation of this xylanase
and uses thereof

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE

STREET: 2000 K St., N.W., Suite 200

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/909,207

FILING DATE: 19-Jul-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/470,953

FILING DATE: 06-JUNE-1995

ATTORNEY/AGENT INFORMATION:

NAME: Wilhem F. Gadiano, Esq.

REGISTRATION NUMBER: 37,136

REFERENCE/DOCKET NUMBER: 4121-40

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-429-0625

TELEFAX: (202) 293-1850

TELEX: 650 383-5605
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 663 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 STRAIN: Bacillus
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 US-09-909-207-1

Query Match 100.0%; Score 663; DB 10; Length 663;
 Best Local Similarity 100.0%; Pred. No. 3.1e-186;
 Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAAATCGTACCGGACAAATTCATTCGACCAACGATGGCTATGATTAATTTGGAAA 60
DB 1 CAAATCGTACCGGACAAATTCATTCGACCAACGATGGCTATGATTAATTTGGAAA 60
QY 61 GATACGGTGGCTCTGGGACAAATGATTCATTCATGCGGTACGTTCCAGTCCCAATGG 120
DB 61 GATACGGTGGCTCTGGGACAAATGATTCATTCATGCGGTACGTTCCAGTCCCAATGG 120
QY 121 AACAACTTTAACACATATTTATTCGTAAGGTAAGTAAATTAATGAAGAACACAAACAC 180
DB 121 AACAACTTTAACACATATTTATTCGTAAGGTAAGTAAATTAATGAAGAACACAAACAC 180
QY 181 CAACAGTTGTGTAAATGTCATTAATCTAAGAGGCACTTCCAAACCAATGGTATGG 240
DB 181 CAACAGTTGTGTAAATGTCATTAATCTAAGAGGCACTTCCAAACCAATGGTATGG 240
QY 241 TATTTATGCGTCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 300
DB 241 TATTTATGCGTCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 300
QY 301 TGGGGCACTGGCGTTCACAGAGCAAGGCTTAAGGGGACCATCATCTGTTATGAGAGA 360
DB 301 TGGGGCACTGGCGTTCACAGAGCAAGGCTTAAGGGGACCATCATCTGTTATGAGAGA 360
QY 361 ACATATGATATCTAGAGAGCTTTAGAGTCAATCAACCTCCATTAAGGGGATGGCACA 420
DB 361 ACATATGATATCTAGAGAGCTTTAGAGTCAATCAACCTCCATTAAGGGGATGGCACA 420
QY 421 TTTAAACAATATGAGTGTTCGAAGATGAAACGACAGTGGACGATTTCTGTACG 480
DB 421 TTTAAACAATATGAGTGTTCGAAGATGAAACGACAGTGGACGATTTCTGTACG 480
QY 481 AACCACTTAAAGCGTGGGAAAATTAGGATGATATGAGGAAAATGTAAGTGTGG 540
DB 481 AACCACTTAAAGCGTGGGAAAATTAGGATGATATGAGGAAAATGTAAGTGTGG 540
QY 541 CTATCTGTAAGAGGCTATCAAGAGTACGAGAGTCTATGATGATGATGATGATGATG 600
DB 541 CTATCTGTAAGAGGCTATCAAGAGTACGAGAGTCTATGATGATGATGATGATGATG 600
QY 601 ATTAAACGTAACCTCTCTCAACTATTAAGATGAGAGAGATTAATTGGATTAAGAAC 660
DB 601 ATTAAACGTAACCTCTCTCAACTATTAAGATGAGAGAGATTAATTGGATTAAGAAC 660
QY 661 AAT 663
DB 661 AAT 663

```

RESULT 2
 US-09-909-207-2
 Sequence 2, Application US/0909207
 Patent No. US2002015181A1
 GENERAL INFORMATION:
 APPLICANT: ANDREE LAHAYE
 ERIC DE BUYL

PIERRE LEDOUX
 RENE DETROZ
 TITLE OF INVENTION: Xylanase, microorganisms produced it,
 DNA molecule, processes for preparation of this xylanase
 and uses thereof
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: WILHELM BRINKS HOFER GILSON & LIONE
 STREET: 2000 K St., N.W., Suite 200
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20006

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/909,207
 FILING DATE: 19-Jul-2001
 CLASSIFICATION: <Unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/470,953
 FILING DATE: 06-JUNE-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Wilhem F. Gadiano, Esq.
 REGISTRATION NUMBER: 37,136
 REFERENCE/DOCKET NUMBER: 4121-40
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-428-0625
 TELEFAX: (202) 293-1850
 TELEX: 650 383-5605

INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 663 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 STRAIN: Bacillus
 FEATURE:
 NAME/KEY: mat. peptide
 LOCATION: 1..663
 NAME/KEY: CDS
 LOCATION: 1..663
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-909-207-2

Query Match 100.0%; Score 663; DB 10; Length 663;
 Best Local Similarity 100.0%; Pred. No. 3.1e-186;
 Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAAATCGTACCGGACAAATTCATTCGACCAACGATGGCTATGATTAATTTGGAAA 60
DB 1 CAAATCGTACCGGACAAATTCATTCGACCAACGATGGCTATGATTAATTTGGAAA 60
QY 61 GATACGGTGGCTCTGGGACAAATGATTCATTCATGCGGTACGTTCCAGTCCCAATGG 120
DB 61 GATACGGTGGCTCTGGGACAAATGATTCATTCATGCGGTACGTTCCAGTCCCAATGG 120
QY 121 AACAACTTTAACACATATTTATTCGTAAGGTAAGTAAATTAATGAAGAACACAAACAC 180
DB 121 AACAACTTTAACACATATTTATTCGTAAGGTAAGTAAATTAATGAAGAACACAAACAC 180
QY 181 CAACAGTTGTGTAAATGTCATTAATCTAAGAGGCACTTCCAAACCAATGGTATGG 240
DB 181 CAACAGTTGTGTAAATGTCATTAATCTAAGAGGCACTTCCAAACCAATGGTATGG 240
QY 241 TATTTATGCGTCTATGTTGATGATGATGATGATGATGATGATGATGATGATGATG 300

```


Db 241 TATTATCGCTCTATGGTTGACTGTTGACCCCTCTTGTGCAATATATATATGTCGACAGT 300
Qy 301 TGGGGCAACTCGCGTCCACGAGGACCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360
Db 301 TGGGGCAACTCGCGTCCACGAGGACCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360
Qy 361 ACATATGATATCTAGGAGACTCTTAGAGTCAATCAACCCCTCATTAAGGGGATTTGCCACA 420
Db 361 ACATATGATATCTAGGAGACTCTTAGAGTCAATCAACCCCTCATTAAGGGGATTTGCCACA 420
Qy 421 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTGAGC 480
Db 421 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTGAGC 480
Qy 481 AACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTGGCG 540
Db 481 AACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTGGCG 540
Qy 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600
Db 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTGTCTAATGTATATAGCAATACACTAAGA 600
Qy 601 ATTAACGGTAACCCCTCTCTCAACTATTAGTAAATGACGAGACATAAATTTTCGATAAAAC 660
Db 601 ATTAACGGTAACCCCTCTCTCAACTATTAGTAAATGACGAGACATAAATTTTCGATAAAAC 660
Qy 661 AAT 663
Db 661 AAT 663

RESULT 3

US-09-909-207-4
; Sequence 4, Application US/09909207
; Patent No. US20020115181A1
; GENERAL INFORMATION:
; APPLICANT: ANDREE LAHAYE
; ERIC DE BUYL
; PIERRE LEDOUX
; RENE DETROZ
; TITLE OF INVENTION: Xylanase, microorganisms produced it,
; and uses thereof
; DNA molecule, processes for preparation of this xylanase
; and uses thereof
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE
; STREET: 2000 K St., N.W., Suite 200
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/909,207
; FILING DATE: 19-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/470,953
; FILING DATE: 06-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilhem P. Gadiano, Esq.
; REGISTRATION NUMBER: 37,136
; REFERENCE/DOCKET NUMBER: 4121-40
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-0625
; TELEFAX: (202) 293-1850
; TELEX: 650 383-5605
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:

LENGTH: 744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
STRAIN: Bacillus
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-909-207-4
Query Match 100.0%; Score 663; DB 10; Length 744;
Best Local Similarity 100.0%; Pred. No. 3.3e-186; Indels 0; Gaps 0;
Matches 663; Conservative 0; Mismatches 0;
Qy 1 CAATTCGTCCACGACAAATTCATTTGGCAACACCATGGCTATGATTAATTTTGGAAA 60
Db 82 CAATTCGTCCACGACAAATTCATTTGGCAACACCATGGCTATGATTAATTTTGGAAA 141
Qy 61 GATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATCGCGGTACGTTCAAGTCCCAATGG 120
Db 142 GATAGCGGTGGCTCTGGGCAATGATTTCTCAATCATCGCGGTACGTTCAAGTCCCAATGG 201
Qy 121 AACAAATGTTAACACATATTATTCCGTAAAGGTAAATAATTCATGAAACACAAACACAC 180
Db 202 AACAAATGTTAACACATATTATTCCGTAAAGGTAAATAATTCATGAAACACAAACACAC 261
Qy 181 CAACAAGTTGGTAACATGTCCATAAACTACGGAGCCAACTTCCAACCAAAATGTAATGCG 240
Db 262 CAACAAGTTGGTAACATGTCCATAAACTACGGAGCCAACTTCCAACCAAAATGTAATGCG 321
Qy 241 TATTTATCGCTCTATGTTGGACCTGTTGACCCCTCTTGTGCAATATTAATTTGTGACAGT 300
Db 322 TATTTATCGCTCTATGTTGGACCTGTTGACCCCTCTTGTGCAATATTAATTTGTGACAGT 381
Qy 301 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 360
Db 382 TGGGGCAACTGGCGTCCACGAGGACCAACGCTTAAGGGGACCATCACTGTTGATGGAGGA 441
Qy 361 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTTGCCACA 420
Db 442 ACATATGATATCTACGAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTTGCCACA 501
Qy 421 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTGACG 480
Db 502 TTTAAACAATATTGGAGTGTTCGAAGATCGAAACGACGAGTGGCAGCATTTCTGTGACG 561
Qy 481 AACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTGGCG 540
Db 562 AACCACTTTAGAGCGTGGGAAACCTTAGGGATGAATATGGGAAATGTATGAAGTGGCG 621
Qy 541 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTAATGTATATAGCAATACACTAAGA 600
Db 622 CTTACTGTAGAGGCTATCAAGTAGCGGAAGTCTAATGTATATAGCAATACACTAAGA 681
Qy 601 ATTAACGGTAACCCCTCTCTCAACTATTAGTAAATGACGAGACATAAATTTTCGATAAAAC 660
Db 682 ATTAACGGTAACCCCTCTCTCAACTATTAGTAAATGACGAGACATAAATTTTCGATAAAAC 741
Qy 661 AAT 663
Db 742 AAT 744

RESULT 4

US-09-909-207-5
; Sequence 5, Application US/09909207
; Patent No. US20020115181A1
; GENERAL INFORMATION:
; APPLICANT: ANDREE LAHAYE
; ERIC DE BUYL
; PIERRE LEDOUX
; RENE DETROZ
; TITLE OF INVENTION: Xylanase, microorganisms produced it,

DNA molecule, processes for preparation of this xylanase
and uses thereof

NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE
STREET: 2000 K St., N.W., Suite 200
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/909,207
FILING DATE: 19-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/470,953
FILING DATE: 06-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Wilhelm F. Gadiano, Esq.
REGISTRATION NUMBER: 37,136
REFERENCE/DOCKET NUMBER: 4121-40
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-0625
TELEFAX: (202) 293-1850
TELEX: 650 383-5605

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 744 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
STRAIN: Bacillus
FEATURE:
NAME/KEY: CDS
LOCATION: 1..744
NAME/KEY: mat_peptide
LOCATION: 82..744
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 1..81
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-909-207-5

Query Match 100.0%; Score 663; DB 10; Length 744;
Best Local Similarity 100.0%; Pred. No. 3.3e-186;
Matches 663; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAAATGTCACGACCAATTCATTCATGGCAACACAGAGGCTATGATTTGGAAA 60
DB 82 CAAATGTCACGACCAATTCATTCATGGCAACACAGAGGCTATGATTTGGAAA 141
QY 61 GATAGGGGCTGCTGGGCAATGATTCATTCATGCGGCTGCTGAGTCCCAATGG 120
DB 142 GATAGGGGCTGCTGGGCAATGATTCATTCATGCGGCTGCTGAGTCCCAATGG 201
QY 121 AACATGTTAAACAATATATTCGTTAAAGTTAAATAATTCATGAAACACAAACAC 180
DB 202 AACATGTTAAACAATATATTCGTTAAAGTTAAATAATTCATGAAACACAAACAC 261
QY 181 CAACAAGTTGTTACATGTCATTAACATGAGGCACTTCCAAACCAATGTTATGGG 240
DB 262 CAACAAGTTGTTACATGTCATTAACATGAGGCACTTCCAAACCAATGTTATGGG 321
QY 241 TATTTATGGGTGATGTTGAGTGTGAGCCCTCTTGGAATATATTTGCGCACT 300

DB 322 TATTTATGGGTGATGTTGAGTGTGAGCCCTCTTGGAATATATTTGCGCACT 381
QY 301 TGGGGCAATGCGCTGTCACAGAGCAACGCTTAAGGGAGCACTGTTGATGAGGA 360
DB 382 TGGGGCAATGCGCTGTCACAGAGCAACGCTTAAGGGAGCACTGTTGATGAGGA 441
QY 361 ACATATGATATCTACAGACTCTTGAAGTCAATCAACCTCCATTAAAGGGATGCCA 420
DB 442 ACATATGATATCTACAGACTCTTGAAGTCAATCAACCTCCATTAAAGGGATGCCA 501
QY 421 TTTAAACAATATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 480
DB 502 TTTAAACAATATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 561
QY 481 AACCACTTTAGAGCGTGGGAAAACTTGAAGTGAATATGAGGAAAAATGATGAGTCCG 540
DB 562 AACCACTTTAGAGCGTGGGAAAACTTGAAGTGAATATGAGGAAAAATGATGAGTCCG 621
QY 541 CTACTGTGAAAGGCTATCAAAAGTGGGAAAGTCTAATGTATATGCAATACCTAAGA 600
DB 622 CTACTGTGAAAGGCTATCAAAAGTGGGAAAGTCTAATGTATATGCAATACCTAAGA 681
QY 601 ATTAAGGTAACCTCTCTCAACTATTAATGATGAGAGCATPACTTGGATTAAC 660
DB 682 ATTAAGGTAACCTCTCTCAACTATTAATGATGAGAGCATPACTTGGATTAAC 741
QY 661 AAT 663
DB 742 AAT 744

RESULT 5
US-09-909-207-10
Sequence 10, Application US/09909207
Patent No. US20020115181A1
GENERAL INFORMATION:
APPLICANT: ANDRE LAMAYE
ERIC DE BUYL
PIERRE LEDOUX
RENE DETROZ
TITLE OF INVENTION: Xylanase, microorganisms produced by it,
DNA molecule, processes for preparation of this xylanase
and uses thereof

NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: WILLIAM BRINKS HOFER GILSON & LIONE
STREET: 2000 K St., N.W., Suite 200
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/909,207
FILING DATE: 19-Jul-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/470,953
FILING DATE: 06-JUNE-1995
ATTORNEY/AGENT INFORMATION:
NAME: Wilhelm F. Gadiano, Esq.
REGISTRATION NUMBER: 37,136
REFERENCE/DOCKET NUMBER: 4121-40
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-0625
TELEFAX: (202) 293-1850
TELEX: 650 383-5605
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:

Db 941 TATTATGCTGTATGTTGAGTGTGACCTCTTGTGCAATATATATGTCGACGT 1000
 Qy 301 TGGGGCAACTGGCGCTCCACAGAGCAACGCTTAAGGGGACCATCATCTGTTGATGAGGA 360
 Db 1001 TGGGGCAACTGGCGCTCCACAGAGCAACGCTTAAGGGGACCATCATCTGTTGATGAGGA 1060
 Qy 361 ACATATGATATCTAGAGACTCTTATGATGATCAACCTCCATTAAGGGGATTCGCACA 420
 Db 1061 ACATATGATATCTAGAGACTCTTATGATGATCAACCTCCATTAAGGGGATTCGCACA 1120
 Qy 421 TTTAAACAATATGAGGTTCGAAAGATGAAACGACGAGTGGACGATTTCTGTCCAC 480
 Db 1121 TTTAAACAATATGAGGTTCGAAAGATGAAACGACGAGTGGACGATTTCTGTCCAC 1180
 Qy 481 AACCACTTAGACGCTGGGAAAATTAGGATGAATATGAGGAAAATGATGAAAGTCGCG 540
 Db 1181 AACCACTTAGACGCTGGGAAAATTAGGATGAATATGAGGAAAATGATGAAAGTCGCG 1240
 Qy 541 CTTACTGTGAAGGCTATCAAAAGTAGCGGAGTGTCTATATGCAATACACTAAGA 600
 Db 1241 CTTACTGTGAAGGCTATCAAAAGTAGCGGAGTGTCTATATGCAATACACTAAGA 1300
 Qy 601 ATTAACGGTAAACCTCTCTCACTATATAGTAATGACGAGACATACTTGGATAAAC 660
 Db 1301 ATTAACGGTAAACCTCTCTCACTATATAGTAATGACGAGACATACTTGGATAAAC 1360
 Qy 661 AAT 663
 Db 1361 AAT 1363
 RESULT 7
 US-09-770-621-1
 ; Sequence 1, Application US/09770621
 ; Patent No. US20010024815A1
 ; GENERAL INFORMATION:
 ; APPLICANT: M ncy1, Arja
 ; APPLICANT: Venmaaper, Jari
 ; APPLICANT: Fagerstr m, Richard
 ; APPLICANT: Lantto, Ralja
 ; APPLICANT: Palohelmo, Marja
 ; APPLICANT: Suominen, Pirkko
 ; APPLICANT: Lahtinen, Tarja
 ; TITLE OF INVENTION: Production and Secretion of Proteins of
 ; NUMBER OF SEQUENCES: 39
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
 ; STREET: 1100 New York Ave., N.W. Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/770,621
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/590,563
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/332,412
 ; FILING DATE: 31-OCT-1994
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/282,001
 ; FILING DATE: 29-JUL-1994
 ; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
 ; NAME: Bugalsky, Lawrence B.
 ; REGISTRATION NUMBER: 35,086
 ; REFERENCE/DOCKET NUMBER: 1050, 0340003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-371-2600
 ; TELEFAX: 202-371-2540
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1375 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: No. US20010024815A1 Relevant
 ; TOPOLOGY: linear (genomic)
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 303..1334
 ; US-09-770-621-1
 Query Match 18.0%; Score 119.4; DB 9; Length 1375;
 Best Local Similarity 57.1%; Pred. No. 8.2e-25;
 Matches 238; Conservative 0; Mismatches 176; Indels 3; Gaps 1;
 Qy 198 GTCCATAAAGTACGAGCACTTCCACCAAAATGTATGCTATATGCTATAG 257
 Db 608 GACCGTACCTTCAACGCTCTTCAACCCGTCGGTAAAGCTACCTACGCTTACGG 667
 Qy 258 TTGACGTGTGACCTCTTGTGCAATATATATATGTCAGAGTTGGGGCAACTGGCGTCC 317
 Db 668 CTGACACGAAACCGCGTGTGAGTACTACGTGTGAGAGCTGGGGCACTTCCGCGC 727
 Qy 318 ACCAGAGCAACGCTTAAGGGGACCATCACTGTTGATGAGGAACTATGATTTACGA 377
 Db 728 CACCGG---CACCTTAAAGGGGACCGTCACACCGACGGGGAACGTACGACATCTAAGA 784
 Qy 378 GACCTTAGAGTCAATCAACCCCTCAATTAAGGGGATTCGCACATTTAAACAATATTGAG 437
 Db 785 GACCTGCGGTACCAACGCGCGTCTCATGAGGACACCGGACCTTCCAGAGTTTGGAG 844
 Qy 438 TGTTCGAAGATCGAAACGACGAGTGGGACGATTTCTGTGCAACCACTTTAGAGCGTG 497
 Db 845 CGTCCGAGGACGAAAGCGGACGAGGACCATCATCGGCAACCACTTTCGACGCGCTG 904
 Qy 498 GAAAACTTAGGAGATGATATGCGGAAATATGATGAATCCGCTTACTGTGAGAGCTTA 557
 Db 905 GCGCCGCGCGCATGAACCTGGGACGACGACATCAATCATGAGGAGCGGCTA 964
 Qy 558 TCAAGTACGCGAAGTGTATATGATATGACATCACTAAGAAATTAACGGTAAACC 614
 Db 965 CCAAGACGCGGTACCTCAACCGTCTCATCACTACGAGGAGTGCAACCCCGGCAACCC 1021
 RESULT 8
 US-10-286-993-1
 ; Sequence 1, Application US/10286993
 ; Publication No. US20030148453A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mantyla, Arja
 ; APPLICANT: Palohelmo, Marja
 ; APPLICANT: Venmaaper, Jari
 ; APPLICANT: Lantto, Ralja
 ; APPLICANT: Fagerstrom, Richard
 ; APPLICANT: Lahtinen, Tarja
 ; APPLICANT: Suominen, Pirkko
 ; APPLICANT: Venmaaper, Jari
 ; TITLE OF INVENTION: Production and Secretion of Proteins in Filamentous
 ; FILE REFERENCE: 1716, 0340004
 ; CURRENT APPLICATION NUMBER: US/10/286,993
 ; FILING DATE: 2002-08-13
 ; PRIOR APPLICATION NUMBER: US/09/120,804
 ; FILING DATE: 1998-07-23
 ; PRIOR APPLICATION NUMBER: PCT/FI97/00037
 ; FILING DATE: 1997-01-24


```

QY 359 GAACATGATATCTAGAGACTCTTGAAGTCATCAACCTTCATTAAAGGAGATTGCCA 418
DB 437 CCACCTAGACATCTATGATGACACCAACAGAGCTTCGATCGCGCACGGCCA 496
QY 419 CATTTAACAAATTTGAGTGTTCGAAGATCCAAACGACGATGGCAGATTTCTGTCA 478
DB 497 CCTTCAACCAATATCTGCTCATCCGCAAAACAGGATTCAGCGGACAGTCACACCG 556
QY 479 GCAACCACTTTAGAGCGTGGGAAAACCTTAGGATGATATGCGGAAAATATGTAAGTCG 538
DB 557 CGAATCACTTCAAGGCTGGGCTAGTCTGGGATGAACTGGGTACCCATTAATCTATCAGA 616
QY 539 CGCTTACTGTAGAGGCTTATCAAGTAGCGGAA 571
DB 617 TTGTTTCCACTGAGGATATGAGAGCGGTA 649

```

RESULT 11
US-10-213-990-70

Sequence 70, Application US/10213990
Publication No. US20030082595A1

GENERAL INFORMATION:

APPLICANT: Jiang, Bo
APPLICANT: Bussey, Howard
APPLICANT: Storms, Reg
APPLICANT: Roemer, Terry
TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
FILE REFERENCE: 10182-019-999
CURRENT APPLICATION NUMBER: US/10/213,990
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 70
LENGTH: 1002
TYPE: DNA
ORGANISM: Aspergillus
US-10-213-990-70

Query Match 16.6%; Score 110; DB 14; Length 1002;
Best Local Similarity 57.8%; Pred. No. 4,3e-22;
Matches 216; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

```

QY 201 CATTAATCTAGGAGCCCACTTCCAAACCAATGTAATGCGTATTATGCGTCTATGTTG 260
DB 336 CATTAACCTTCTGCGAGCTTCAATCTTCGGAATGCTTACCTGCTGATGATGATG 395
QY 261 GACTGTGACCTCTTGTGGAATATATATGTCGACATGGGGA---ACTGGGCTCC 317
DB 396 GACTACCAACCCCTAGTGAATATCTAATCTCTCGAAGCTATGGAATTAACATCTCG 455
QY 318 ACCAGAGCAAGCGCTTAGGGGACATCACTGTATGAGGAAATATGATATCTACGA 377
DB 456 CTCGGGATGAGCACAAGGACCTCACCAGCATGATCACCCTAGACATCTATATA 515
QY 378 GACTTGAAGTCAATCAACCTTCATTAGGGGATGCGACATTTAAACATATTGAG 437
DB 516 GCACCAACAGGTCAACAGCCTTCATCGTCGGCAGGCACTTCAACCAATATGTCG 575
QY 438 TGTTCGAAGATGAAACGACAGATGTCAGATTTTGTTCAGAACCACTTTAGAGCGTG 497
DB 576 CATCCGCCAAACMAAGCATCCAGCGGCAAGTCACACCGGMACTTCAAGGCTCG 635
QY 498 GGAACCTTAGGGATGATATGGGAAAATGTATGAAGTCGCGCTTACTGTAGAAAGCTA 557
DB 636 GCGTACTCTGGGGATTAACCTGGGTACCCATTAATCTCAAGATTGTTTCACTGAGGGATA 695
QY 558 TCAAGTAGCGGAA 571
DB 696 TGAGAGCAAGCGGTA 709

```

RESULT 12

US-10-213-990-68
Sequence 68, Application US/10213990
Publication No. US20030082595A1

GENERAL INFORMATION:

APPLICANT: Jiang, Bo
APPLICANT: Bussey, Howard
APPLICANT: Storms, Reg
APPLICANT: Roemer, Terry
TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUMIGATUS ENCODING INDUSTRIAL
FILE REFERENCE: 10182-019-999
CURRENT APPLICATION NUMBER: US/10/213,990
NUMBER OF SEQ ID NOS: 72
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 68
LENGTH: 705
TYPE: DNA
ORGANISM: Aspergillus
FEATURE:
NAME/KEY: CDS
LOCATION: (1) ... (705)
US-10-213-990-68

Query Match 16.4%; Score 108.8; DB 14; Length 705;
Best Local Similarity 51.9%; Pred. No. 8e-22;
Matches 298; Conservative 0; Mismatches 267; Indels 9; Gaps 2;

```

QY 6 CGTACCGCAATTCATTTGGCAACACAGATGGGATATATGATTTGGAAGATAG 65
DB 120 CGGACCCCAAGCTCCACCGGTGGAACAGGCTACTACTCTCTCTGACTGATGG 179
QY 66 CGGTGCTCTGGGACATATGATTTCTCAATCATGCGGATAGCTTCAAGTCCCATGACAA 125
DB 180 CGGCGGCAACGTGACTACACCAATGCGCGGTGCTGTACTCCGTCACTGAGGAA 239
QY 126 TGTTAACAATATTTATTCCTTAAGTAAATTTCAATGAACAAACACCAACA 185
DB 240 CGTGGCAACTTTTGTGCGTGAAGGAGC-----TGAAACCTGGAGCGGTAGGTACG 293
QY 186 AGTGTAACTATGTCATTAATCTAGGAGCAACCTTCCAAACCAATGTAATGCGTATT 245
DB 294 AGCTTAAGTGAACATCACTACAGGAGGAGCTTCAACCCACGCGCAATGCTACT 353
QY 246 ATGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 305
DB 354 GCGTGTCTACGCTGGACCAACCAACCCCTTATGATGATGATGATGATGATG 413
QY 306 CA---ACTGGGCTCCACAGAGCAAGCCTTAAGGGGACATCACTGTTGATGAGAAC 362
DB 414 TACATACAAACCCCGACAGCGGCTTCACTTCAAGGGGCACTGTCAACCGAGGTGGAC 473
QY 363 ATATGATATCTACAGAGCTCTTAAGTCAATCAACCTTCATTAAAGGATTTGCCAAT 422
DB 474 TTACAACATCTACAGCGCGCTTCCTTCAATAGCTCTCCATCGAAGCACCAAGACTT 533
QY 423 TAAACATATGAGTGTGTCGAAGATCGAAACGACAGATGCGACGATTTCTGACGAA 482
DB 534 CACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 593
QY 483 CCACTTGAAGCGTGGGAAAACCTTAGGATGATATGCGGAAAATATGTAAGTCGCGCT 542
DB 594 CCACTTGAAGCGTGGGAAAACCTTAGGATGATATCACTTGGGAACTCAACATCAAGATTGT 653
QY 543 TACTGTAGAGGCTATCAAGTAGCGGAGTCT 576
DB 654 CGCAGTGAAGGATTAACAGAGCGGATCTGCT 687

```

RESULT 13

US-10-299-393-1

Sequence 1, Application US/10299393

```
; Publication No. US20030108642A1
; GENERAL INFORMATION:
; APPLICANT: Sabatier, Alain
; APPLICANT: Fish, Neville Marshall
; APPLICANT: Haigh, Nigel Paterson
; TITLE OF INVENTION: PENICILLIUM FUNICULOSUM STRAIN USEFUL
; TITLE OF INVENTION: FOR THE PRODUCTION OF ENZYMES
; FILE REFERENCE: A32917-PCT-USA-1 (072667.0183)
; CURRENT APPLICATION NUMBER: US/10/299,393
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 09/462,246
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: EPO 9801161.5
; PRIOR FILING DATE: 1998-05-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2898
; TYPE: DNA
; ORGANISM: Penicillium funiculosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1317)...(1589)
; FEATURE:
; NAME/KEY: Intron
; LOCATION: (1590)...(1642)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1644)...(2042)
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (570)...(576)
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (724)...(730)
; US-10-299-393-1

Query Match 16.0%; Score 106; DB 14; Length 2898;
Best Local Similarity 56.5%; Pred. No. 1.1e-20;
Matches 218; Conservative 0; Mismatches 165; Indels 3; Gaps 1;

QY 195 CATGTCCTAATACTAGGAGCCAACTTCCACCAAAATGGTATGGTATTATGGCTCTA 254
Db 1640 CAGGACTGTACGTACTCAGGAGAAATTAACCCCTCTGGAAACGCTTATTTGGCTGTCTA 1699
QY 255 TGGTTGGAGCTGTGACCCCTCTTGTGCAATATTATATTGTGCACTTGGGCAACTGGCG 314
Db 1700 CGGTGGACACAGATCTCTTGTGCAATATTATATTGTGCACTTGGAGTCTTACGGCACTATAA 1759
QY 315 TCCACCA---GGAGCAACGCCCTTAAGGGACCATCACTGTGATGGAGGAACATATGATAT 371
Db 1760 CCCATCATCTGGACTTACTTCACTTTGGCCAGGTCACTAGCGATGTTGGCACTAGGATAT 1819
QY 372 CTACAGACTCTTAGAGTCAATCAACCCCTCCATTAAGGGAGTTGCCAATTTAAACAATA 431
Db 1820 CTACTCAACCCACGCGTGTCAACAGCCTTCCATTTGAGGAACTTCCACCTTCAACAGTA 1879
QY 432 TTGGAGTGTTCGAAGATCGAAACGCAACGAGTGGCAGGATTTCTGTCAGCAACCACTTTAG 491
Db 1880 CTGGTCACTTTCACCGAAGAGCGAGTGGCGGAACTGTCCACCGCAACCACTTTGC 1939
QY 492 AGCGTGGGAAAATTAGGAGTGAATATATGGGAAAATGTATGAAGTCGGCGCTTACTGTAGA 551
Db 1940 AGCATGGAAGGCACCTTGGACTTGAATATGGGCACTTATAACTATATGATGTTGTGCCACGA 1999
QY 552 AGGCTCAAGTACGGAAGTGCTA 577
Db 2000 AGGCTACGAGACGAGTGGCTCTAGTA 2025

RESULT 14
US-10-213-990-64
; Sequence 64, Application US/10213990
; GENERAL INFORMATION:
; APPLICANT: Bussey, Howard
; APPLICANT: Stormer, Terry
; APPLICANT: Rosmer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUNIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Aspergillus
; FEATURE:

; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Stormer, Reg
; APPLICANT: Rosmer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUNIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 712
; TYPE: DNA
; ORGANISM: Aspergillus
; US-10-213-990-64

Query Match 15.8%; Score 104.8; DB 14; Length 712;
Best Local Similarity 55.8%; Pred. No. 1.2e-20;
Matches 221; Conservative 0; Mismatches 172; Indels 3; Gaps 1;

QY 189 TGGTAACATGTCCATAAATACGAGGCAACTTCCAAACCAATGGTATGGTATTTATG 248
Db 304 TGGATTTCAGAGCGGTCCACTACAGCGGCTCCTGGCAGACCAAGCGGTAACCTCTC 363
QY 249 CGTCTATGTTGGACTGTTGACCCCTCTTGTGCAATATTATATTTCGACACGTTGGGCAA 308
Db 364 CGTGTACGGCTGGAGACCAAGTCCGCTTGTGCAATTTACATCGTGGAGAGTTACGGCTC 423
QY 309 CTGGCGTCCACCA---GGAGCAACGCCCTTAAGGGACCATCACTGTGATGGAGGAACATA 365
Db 424 CTATGACCCCTCCACGGGAGCCACCCATCTCGGCACCGTCGAGAGCGAGGGCCACGTA 483
QY 366 TGATATCTACGAGACTCTTAGAGTCAATCAACCTCCATTAAGGGAGTTGCCAATTTAA 425
Db 484 CAACCTCTTCAAGACGACGCGGAGCAATGCGCCGCTCCATCCAGGGCAGCGCTACTTTGA 543
QY 426 ACAATATTGGAGTGTTCGAGATCGAAACGCAACGACGAGTGGCACGATTTCTGTCAGCAACCA 485
Db 544 CCAGTACTGTGTCGGTTCGGACTTCGCACCGCAGAGTGGAACTGTGACGAGGAAGAACCA 603
QY 486 CTTTAGACGCTGGGAAAATCTTAGGGATGAATATATGGGAAAATGTATGAAGTCGCGCTTAC 545
Db 604 CTTTGATGCGTGGAGAAATGCGGCTCTGCAATTCGGGAACTTTGACTATATGATGTTGTC 663
QY 546 TGTAGAAGGCTATCAAGTAGTCGGAAGTGTAAATGT 581
Db 664 GACGAGGGGTACCAAGACGACGCGCTCTGCTACTAT 699

RESULT 15
US-10-213-990-65
; Sequence 65, Application US/10213990
; Publication No. US20030082595A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Bussey, Howard
; APPLICANT: Stormer, Reg
; APPLICANT: Rosmer, Terry
; TITLE OF INVENTION: NUCLEIC ACIDS OF ASPERGILLUS FUNIGATUS ENCODING INDUSTRIAL
; FILE REFERENCE: 10182-019-999
; CURRENT APPLICATION NUMBER: US/10/213,990
; CURRENT FILING DATE: 2002-08-05
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Aspergillus
; FEATURE:
```

/ NAME/KEY: CDS
; LOCATION: (1)...(666)
US-10-213-990-65

Query Match 15 6%; Score 103.6; DB 14; Length 666;
Best Local Similarity 56.3%; Pred. No. 2.7e-20;
Matches 215; Conservative 0; Mismatches 164; Indels 3; Gaps 1;

Qy	203 TAACTACGAGCCCACTTCAACCAATGTATGCGTATTATGCGTCTATGTTGA	262
Db	272 TCACCTACAGCGGCTCTCGACAGCAAGGGAACGGCTACCTCTCGTACGGCTGA	331
Qy	263 CTGTTGACCTCTTGTGAAATTTATATTGTGACAGTTGGGCAACTGGCTCCACA	321
Db	332 CGACCAATCGGCTGTGAAATTTCTACATGTTGAGAGTTACGGCTCTCTATGACCCCTCCA	391
Qy	322 --GGAGCAACGGCTAAGGGGACCATGACTGTTGATGAGAGAACTATGATCTACGAGA	379
Db	392 CGGAGCCACCACTCTGSCACCGTCGAGAGCGACGGGCCACGTACAACTCTTACAGAA	451
Qy	380 CTCTTAGAGTCAATCAACCCCTCCATTAGGGGATTCACATTTAAACAATATTGGAGTG	439
Db	452 CGACCGGAGCAATGGCGCGCTCATCCAGGGCACGGCTACTTTTGACAGTACTGGTCGG	511
Qy	440 TTGGAAGATCGAAACGACGAGTGGCAGATTCTGTCAACAACCACTTTAGAGCGTGG	499
Db	512 TTCGGACTTCGCAACCGGACAGTGGAACTGTGACGACGAACAACCACTTTGATGCGTGA	571
Qy	500 AAACTTAGGATGAAATATGGGAAATGATGAAGTCGCGCTTACTGTAGAAGCTATC	559
Db	572 GAAATCGGGTCTGCAATTGGGAACTTTGACTATATGATGTTGCGACGAGGGGTACC	631
Qy	560 AAAGTAGCGAAGTCTAATGT	581
Db	632 AGAGCAGCGGCTCTGCTACTAT	653

Search completed: October 20, 2003, 13:35:44
Job time : 1600 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 20, 2003, 10:40:41 ; Search time 2045 Seconds
(without alignments)
7879.642 Million cell updates/sec

Title: US-09-909-207-1

Perfect score: 663

Sequence: 1 CAATGTCACCGAATTC.....TAATTTGGATAAAACAAT 663

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86.6	13.1	738	14	CD464145
2	83.4	12.6	921	14	CD458837
C 3	82.8	12.5	603	28	AQ160513
C 4	82.2	12.4	670	28	AQ447125
					mgxb0005C

C 5	82.2	12.4	720	28	AQ361561
C 6	82.2	12.4	750	28	AQ160254
C 7	81.6	12.3	786	28	AQ325248
C 8	81.4	12.3	583	28	AQ399120
C 9	76.6	11.6	561	28	AQ396475
C 10	74.6	11.3	753	28	AQ448084
C 11	74.4	11.2	418	28	AQ398756
C 12	69.8	10.5	520	13	BQ165950
C 13	66.2	10.0	571	13	BQ471960
C 14	63.2	9.5	473	14	CD464005
C 15	60.4	9.1	483	13	BQ664593
C 16	53.6	8.1	646	28	AQ449078
C 17	49.4	7.5	490	10	BF200865
C 18	49.4	7.5	532	12	BM134812
C 19	46.6	7.0	503	13	BQ166480
C 20	45.2	6.8	587	14	CD033274
C 21	41.6	6.3	488	12	BM135798
C 22	40.2	6.1	665	10	BB664921
C 23	40.2	6.1	1319	11	AK053390
C 24	39.6	6.0	704	13	BQ110128
C 25	39	5.9	1146	13	BX446722
C 26	38.8	5.9	588	28	AQ630020
C 27	38.8	5.9	750	14	CA447770
C 28	38.8	5.9	1809	11	BC022353
C 29	38.2	5.8	615	28	AZ059161
C 30	38	5.7	540	12	BM897662
C 31	38	5.7	933	29	CNS01TBY
C 32	37.6	5.7	279	10	BB430376
C 33	37.2	5.6	403	28	AQ461126
C 34	37.2	5.6	861	29	CNS0117J
C 35	37	5.6	690	28	BZ023864
C 36	37	5.6	702	28	BH938430
C 37	36.8	5.6	646	14	CB160626
C 38	36.8	5.6	1101	29	CNS01616
C 39	36.6	5.5	232	9	AV315991
C 40	36.6	5.5	437	12	BM107570
C 41	36.6	5.5	549	14	N20977
C 42	36.6	5.5	685	28	BH989520
C 43	36.6	5.5	698	29	AG007299
C 44	36.6	5.5	705	29	AG007300
C 45	36.6	5.5	714	28	BZ045202

ALIGNMENTS

RESULT 1
CD464145
LOCUS
DEFINITION
ETH1_48_B06.g1_A002 Ethylene-treated seedlings Sorghum bicolor CDNA
clone ETH1_48_B06_A002 5', mRNA sequence.
CD464145
VERSION
CD464145.1 GI:31385413
KEYWORDS
EST.
SOURCE
Sorghum bicolor (sorghum)
ORGANISM
Sorghum bicolor
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 738)
REFERENCE
Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S., Klein
N.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Salzman,R., Chua Tan
N., Gonzalez,M., Lane,S., Miller,V., Nanda,P., Olaseinde,O.,
Eastman,A. and Pratt,L.H.
AUTHORS
An EST database from Sorghum: 1-aminocyclopropane-1-carboxylic acid
(ACC)-treated seedlings
Unpublished
TITLE
Other ESTs: ETH1_48_B06.b1_A002
JOURNAL
Contact: Cordonnier-Pratt MM
COMMENT
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860

Fax: 706 583 0210
 Email: mmpirat@uga.edu
 Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in the Human Genome Center, University of Tokyo Institute of Medical Science; plant material and RNA prepared at Texas A & M University; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below phred quality 15. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.

Seq primer: 5'GCTCTGCTCAAAAGCTGCG).

FEATURES

source

Location/Qualifiers

1..738
 /organism="Sorghum bicolor"
 /mol_type="rRNA"
 /culturvar="BTx623"
 /db_xref="taxon:4558"
 /clone="ETH1.48.B06.A002"
 /lab_host="DH10B-T1 phage-resistant E. coli"
 /note="Ethylene-creating seedling"
 /note="Vector: pMB18S-FL3; Site 1: XhoI; Site 2: XhoI; The library was prepared from polyA+ RNA from seedlings grown in hydroponic culture. At 8 days of age, medium was supplemented with 0.5 mM 1-aminocyclopropane-1-carboxylic acid (ACC) to induce endogenous ethylene (ETH) production. Roots and shoots were harvested after 27 and 72 hr and material from both time points was combined prior to RNA isolation. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pMB18S-FL3 vector (5'-prime DraIII site is CACGCTGG, 3'-prime DraIII site is CACCATGTG). XhoI excises the cDNA insert."

BASE COUNT

168 a 242 c 164 g 164 t

ORIGIN

Query Match 13.1%; Score 86.6; DB 14; Length 738;

Best Local Similarity 53.5%; Pred. No. 2.7e-09; Mismatches 174; Indels 3; Gaps 1;

Matches 204; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

200 CCATAAATACGAGCACTTCCAAACCAATGTAATCGATTATTCGCTATGTT 259
 345 CCATCAAGTACTCTGTACTTCAACATCAAGCAACGCTACGCTGTTACGAT 404
 260 GGAAGTGTGACCTCTTGTGAATATTATTTGTCAGCACTTGGGCAACTGGCTCCAC 319
 405 GGACTCAGAACCTCTCATCGAGTACTACATCGTTGAGAACTTGGCACTTCAACCCCT 464
 320 CA---GAGCAAGCCTTAGGGGACCATGCTGTGATGAGGAACATATGATCTAGC 376
 465 CTTCCGGCGGCAAGAAAGGGAGAGTCACTGTGACGATCTGTCTACGACATCTACG 524
 377 AGACTTTAGATCAATCAACCTCCATTAAAGGGATTCACATTTAAACAATATTGA 436
 525 TCAGACACCGGTGTCAACGCCCTTCATTGAGGTAAACAAGCCTTTCAGCAGTTGCT 584
 437 GTGTTGAAGATCGAAAGCAGAGTGCACATTTCTGTACGAAACCACTTTAGAGCT 496
 585 CGTTTCAGAGCAACAGCATCCAGCGGATCCGTCACACCGGTGCTCACTTCCAGGCT 644
 497 GGGAAACTTAGAGATGATGATGAGGAAATGTATGAATCGGCTTACGTAGAGAGCT 556
 645 GGAATAAATGTGGCTCAACCTTGTGTAACCAACATGATCTTGTGTAAGGCT 704
 557 ATCAAGTAGCGAAGTGTCTA 577
 705 ACTACAGCTCTGCTCTGCA 725

RESULT 2
 CD458837 921 bp mRNA linear EST 03-JUN-2003
 LOCUS Fg08_04b10_A Fg08_AAFc_ECORC_Fusarium_graminearum_complex_substrate
 DEFINITION Gibberella zeae cDNA clone Fg08_04b10, mRNA sequence.

ACCESSION CD458837
 VERSION CD458837.1 GI:31373577
 KEYWORDS EST.
 SOURCE Gibberella zeae
 ORGANISM Gibberella zeae

REFERENCE
 AUTHORS Hypocotylomycetidae; Hypocotylomycetidae; Sordariomycetes; Eukaryota; Fungi; Ascomycota; Hypocotylomycetidae; Gibberella. 1 (bases 1 to 921)
 Watson, R.J., Heyes, R., Chapados, J., Couroux, P., Harris, L.J., Hartori, J., Lacroix, C., Ouellet, T., Robert, L.S., Singh, J.A., Sprout, D. and Tinker, N.A.

TITLE

A cDNA library prepared from *Fusarium graminearum* grown on a complex plant substrate

JOURNAL

Unpublished
 Contact: Watson, Robert J.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-Food Canada
 Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
 CANADA
 Tel: (613) 759-1655
 Fax: (613) 759-1701
 Email: watsonrj@agr.gc.ca.

FEATURES

source

Location/Qualifiers

1..921
 /organism="Gibberella zeae"
 /mol_type="rRNA"
 /strain="DAOM 180378"
 /db_xref="taxon:5518"
 /clone="Fg08_04b10"
 /tissue="Fg08_04b10"
 /dev_stage="Asexual"
 /lab_host="E. coli DH10B"
 /note="Vector: pBluescript II+; Site 1: EcoRI; Site 2: XhoI; *Fusarium graminearum* grown on a complex plant substrate-- wheat leaves treated to remove most of the low molecular weight, water-soluble components."

BASE COUNT

203 a 254 c 232 g

ORIGIN

Query Match 12.6%; Score 83.4; DB 14; Length 921;

Best Local Similarity 52.4%; Pred. No. 1.4e-08; Mismatches 186; Indels 3; Gaps 1;

Matches 208; Conservative 0; Mismatches 186; Indels 3; Gaps 1;

200 CCATAAATACGAGCACTTCCAAACCAATGTAATCGATTATTCGCTATGTT 259
 363 CCATCAAGTACTCTGTACTTCAACATCAAGCAACGCTACGCTGTTACGAT 422
 260 GGAAGTGTGACCTCTTGTGAATATTATTTGTCAGCACTTGGGCAACTGGCTCCAC 319
 423 GGACTCAGAACCTCTCATCGAGTACTACATCGTTGAGAACTTGGCACTTCAACCCCT 482
 320 CAGGA---GCAAGCCTTAGGGGACCATGCTGTGATGAGGAACATATGATCTACG 376
 483 GAGCAGAGCTGACGACGAGTACCTGTACACCGAGGTACCTACGATCTCTATA 542
 377 AGACTTTAGATCAATCAACCTCCATTAAAGGGATTCACATTTAAACAATATTGA 436
 543 TGTCAACCGGTGTCAACAGCCTTTCATTGAGGTAAACAAGCCTTTCAGCAGTTGCT 602
 437 GTGTTGAAGATCGAAAGCAGAGTGCACATTTCTGTACGAAACCACTTTAGAGCT 496
 603 CCATCCGGCGGCAAGAGGTACACCGGCTCGTCAACATGCAAGAACCACTTCAATGCTT 662
 497 GGGAAACTTAGAGATGATGATGAGGAAATGTATGAATCGGCTTACGTAGAGAGCT 556
 663 GGAGATCTGCTGCTGATGAACCTCGAAGACCACTTACCAAGTTCTGGCTACTGAGGCT 722
 557 ATCAAGTAGCGAAGTGTCTAATGATATGACATAC 593
 723 ACCAAGAGAGTGGCTCATCTTCTATATGTCACAC 759

```

RESULT 3
AQ160513/c
LOCUS
DEFINITION
  mgxb0006C21r CUGI Rice Blast BAC Library Magnaporthe grisea genomic
  clone mgxb0006C21r, genomic survey sequence.
ACCESSION
  AQ160513
VERSION
  AQ160513.1 GI:3557502
KEYWORDS
  GSS.
SOURCE
  Magnaporthe grisea (anamorph: Pyricularia grisea)
  Magnaporthe grisea
  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
  Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
  1 (bases 1 to 603)
  Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
  Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
  A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
  Genome
JOURNAL
  Unpublished
COMMENT
  Contact: Dean RA
  Clemson University Genomics Institute
  100 Jordan Hall, Clemson University, Clemson, SC 29634
  Tel: 864 656 5737
  Fax: 864 656 4293
  Email: rdeane@clemson.edu
  Seq primer: GGAAACAGCTATGACCATG
  Class: BAC ends
  High quality sequence stop: 424.
  Location/Qualifiers
    1..603
      /organism="Magnaporthe grisea"
      /mol_type="genomic DNA"
      /strain="70-15"
      /db_xref="taxon:148305"
      /clone="mgxb0006C21r"
      /tissue_type="Protoplasts"
      /lab_hosts="E. coli DH10B"
      /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
      Rice blast is one of the most devastating fungal diseases
      of rice world wide. It is a filamentous ascomycete with
      a haploid genome (n=7) of approximately 40 Mbp. Rice
      blast is an important model fungal pathogen for studying
      numerous aspects of the fungal-host interaction. In
      order to facilitate genome wide analysis, a BAC library
      containing 9216 clones with an average insert size of 130
      kbp was constructed. This library represents greater
      than 25X genome coverage. High density colony filters
      are available upon request."
      104 a 152 c 202 g 144 t 1 others
BASE COUNT
  104 a 152 c 202 g 144 t 1 others
ORIGIN
  Query Match 12.5%; Score 82.8; DB 28; Length 603;
  Best Local Similarity 52.8%; Pred. NO. 1.9e-08;
  Matches 227; Conservative 0; Mismatches 197; Indels 6; Gaps 2;

Qy 154 AAAAAATCAATGAACAAACACCAACAAAGTTGGTAAATGTCATGTCATAACTACGGA 213
  |||||
Db 444 AAGAAAAAAGAGACTACCAACAAACAAACAAACAGCCGGTCTCACTACTTCG 385

Qy 214 GCCAACTTCCAAACAAATGGTAAATGCGTATTATGCGTCTATGGTTGACCTGTTGACCCCT 273
  |||||
Db 384 GGCAGCTACAGCCCCAGCGCAACTCATACCTGGCGGTCTACGGCTGGAGCGGCAACCG 325

Qy 274 CTTGTCGATATTATATTGTCAGATGTTGGGCAACTCGGCTCCACCA---GGAGCAACG 330
  |||||
Db 324 CTGATCGAGTACTACGTGGTGAGAGCTTTTGCGAGCTACAAACCCGTCGTGGCGGCCACC 265

Qy 331 CCTAAGGGGACCATCAGCTGTTGATGGAGGACATATGATATCTACGAGACTCTTAGAGTC 390
  |||||
Db 264 AACCAGCGGTCTTCACTCGGAGCGGAGCACTACGACATCTCTGGTCAGCACCCGCTAC 205

```

```

Qy 391 AATCAACCTCTCAATTAAGGGGATTGCCACATTTAAACAATATTGGAGTGTTCGAAGATCG 450
  |||||
Db 204 AACCAGCGCTCTCATCGACGCGCACCAAGACCTTTTCAGCAGTGTCTGTCGGTGGCGGCAAC 145

Qy 451 AAACGACGAGTGGCAGCATTTCTGTGAGCAACACATTTAGAGCGTGGGAAACCTTAGGG 510
  |||||
Db 144 AAGCGCGCAGCGGCGCACCGTCACCTTTGCCAACACACGTCAACGGCTGGCGGCAACGCCGCGC 85

Qy 511 ATGAATATGGGAAAAATGTATGAAGTCGCG---CTTACTGTAGAGGCTATCAAAAGTAGC 567
  |||||
Db 84 CTCAACCTCGGCACCACTGGAACTACACAGATCTCGACCGTCGAGGGCTACCAACAGCAGC 25

Qy 568 GGAAGTGTCTA 577
  |||||
Db 24 GGCTCCGCCA 15

```

```

RESULT 4
AQ447125/c
LOCUS
DEFINITION
  mgxb0005C20f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
  clone mgxb0005C20f, genomic survey sequence.
ACCESSION
  AQ447125
VERSION
  AQ447125.1 GI:4576262
KEYWORDS
  GSS.
SOURCE
  Magnaporthe grisea (anamorph: Pyricularia grisea)
  Magnaporthe grisea
  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
  Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
  1 (bases 1 to 670)
  Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
  Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
  A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
  Genome
JOURNAL
  Unpublished
COMMENT
  Contact: Dean RA
  Clemson University Genomics Institute
  100 Jordan Hall, Clemson University, Clemson, SC 29634
  Tel: 864 656 5737
  Fax: 864 656 4293
  Email: rdeane@clemson.edu
  Seq primer: TAATACGACTCACTATAGGG
  Class: BAC ends
  High quality sequence stop: 263.
  Location/Qualifiers
    1..670
      /organism="Magnaporthe grisea"
      /mol_type="genomic DNA"
      /strain="70-15"
      /db_xref="taxon:148305"
      /clone="mgxb0005C20f"
      /tissue_type="Protoplasts"
      /lab_hosts="E. coli DH10B"
      /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
      Rice blast is one of the most devastating fungal diseases
      of rice world wide. It is a filamentous ascomycete with
      a haploid genome (n=7) of approximately 40 Mbp. Rice
      blast is an important model fungal pathogen for studying
      numerous aspects of the fungal-host interaction. In
      order to facilitate genome wide analysis, a BAC library
      containing 9216 clones with an average insert size of 130
      kbp was constructed. This library represents greater
      than 25X genome coverage. High density colony filters
      are available upon request."
      114 a 192 c 198 g 166 t
BASE COUNT
  114 a 192 c 198 g 166 t
ORIGIN
  Query Match 12.4%; Score 82.2; DB 28; Length 670;
  Best Local Similarity 54.6%; Pred. No. 2.6e-08;
  Matches 209; Conservative 0; Mismatches 168; Indels 6; Gaps 2;

```

QY	201	QY	202	QY	203	QY	204	QY	205	QY	206	QY	207	QY	208	QY	209	QY	210	QY	211	QY	212	QY	213	QY	214	QY	215	QY	216	QY	217	QY	218	QY	219	QY	220	QY	221	QY	222	QY	223	QY	224	QY	225	QY	226	QY	227	QY	228	QY	229	QY	230	QY	231	QY	232	QY	233	QY	234	QY	235	QY	236	QY	237	QY	238	QY	239	QY	240	QY	241	QY	242	QY	243	QY	244	QY	245	QY	246	QY	247	QY	248	QY	249	QY	250	QY	251	QY	252	QY	253	QY	254	QY	255	QY	256	QY	257	QY	258	QY	259	QY	260	QY	261	QY	262	QY	263	QY	264	QY	265	QY	266	QY	267	QY	268	QY	269	QY	270	QY	271	QY	272	QY	273	QY	274	QY	275	QY	276	QY	277	QY	278	QY	279	QY	280	QY	281	QY	282	QY	283	QY	284	QY	285	QY	286	QY	287	QY	288	QY	289	QY	290	QY	291	QY	292	QY	293	QY	294	QY	295	QY	296	QY	297	QY	298	QY	299	QY	300	QY	301	QY	302	QY	303	QY	304	QY	305	QY	306	QY	307	QY	308	QY	309	QY	310	QY	311	QY	312	QY	313	QY	314	QY	315	QY	316	QY	317	QY	318	QY	319	QY	320	QY	321	QY	322	QY	323	QY	324	QY	325	QY	326	QY	327	QY	328	QY	329	QY	330	QY	331	QY	332	QY	333	QY	334	QY	335	QY	336	QY	337	QY	338	QY	339	QY	340	QY	341	QY	342	QY	343	QY	344	QY	345	QY	346	QY	347	QY	348	QY	349	QY	350	QY	351	QY	352	QY	353	QY	354	QY	355	QY	356	QY	357	QY	358	QY	359	QY	360	QY	361	QY	362	QY	363	QY	364	QY	365	QY	366	QY	367	QY	368	QY	369	QY	370	QY	371	QY	372	QY	373	QY	374	QY	375	QY	376	QY	377	QY	378	QY	379	QY	380	QY	381	QY	382	QY	383	QY	384	QY	385	QY	386	QY	387	QY	388	QY	389	QY	390	QY	391	QY	392	QY	393	QY	394	QY	395	QY	396	QY	397	QY	398	QY	399	QY	400	QY	401	QY	402	QY	403	QY	404	QY	405	QY	406	QY	407	QY	408	QY	409	QY	410	QY	411	QY	412	QY	413	QY	414	QY	415	QY	416	QY	417	QY	418	QY	419	QY	420	QY	421	QY	422	QY	423	QY	424	QY	425	QY	426	QY	427	QY	428	QY	429	QY	430	QY	431	QY	432	QY	433	QY	434	QY	435	QY	436	QY	437	QY	438	QY	439	QY	440	QY
----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----	-----	----

order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

Query Match	12.4%	Score 82.2	DB 28	Length 720
Best Local Similarity	54.1%	Pred. No. 2.6e-08		
Matches 236	Conservative 0	Mismatches 189	Indels 11	Gaps 3
OY	148	AAAGTAAAAAATTCATGTAAGAAACAAACACACCAACAGTTGGTAAATGTCATTAAC	207	
Db	445	AAAGGAAAAAAGAGACTTAACAACAACAACAACAAAAA-----CAACCGCGTCATCAAC	391	
OY	208	TACGAGACCAACTTCCAAACCAATGATGAATGCGATTTATGCGCTATATGTTGGACTGTT	267	
Db	390	TATCTGGGACGCTACAGCCCNACAGGACCTCATACCTGGCCGCTACAGCGTGGACGCG	331	
OY	268	GACCCCTTTGTCGAATATATATATGTCGACAGTTGGGGCAACTGGCGTCCACCA---GGA	324	
Db	330	AAACCCGCTGATCGAGTACTACGTGTGTGAGAGAGCTTTGGACGTACAAACCCGTCGTGGGC	271	
OY	325	GCAACGCTTAAGGGGACCATCATCTGTTGATGAGAGAAACATATGATATCTACGAGACTCTT	384	
Db	270	GCCACCAACCGCGGGGTCCTTACACTCCGGAAGGCGACCTACGCAATCTGCTCAGACAC	211	
OY	385	AGACTCAATCAACCTCTCATTTAAGGGGATTTGCCATTTAAACAAATTTGAGATGTTGGA	444	
Db	210	CGTACAAACAGCCCTTCATTCAGAGGACCAAGACCTTTAGCACTTCTGTCGTCGCGC	151	
OY	445	AGATCGAAAGCGACAGATGGACAGATTTCTGTGACGAACACATTTAGAGCGTGGGAAAC	504	
Db	150	CGCAACAAGCGCGCGCAAGCGGACCGGTACCTTTGGCCAAACGTCACAGCGCTGGCGCAC	91	
OY	505	TTAGGAAATGAATATGSGGAAATATGATGAGTGCG---CTTACTGTAGAAGCTATCA	561	
Db	90	GCCGCGCTCAACCTCGGCAACCAAGTGAACCTACAGATCTCGCGCTGAGGGCTACAC	31	
OY	562	AGTAGCGGAAGTGTCTA	577	
Db	30	AGCAGCGAGCTCCGCCA	15	

RESULT 6

LOCUS A0160254/c 750 bp DNA linear GSS 09-SEP-1998

DEFINITION mgxb0003l19r CUGI Rice Blast BAC Library Magnaporthe grisea genomic clone mgxb0003l19r, genomic survey sequence.

ACCESSION A0160254

VERSION A0160254.1 GI:357243

KEYWORDS GSS.

SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)

ORGANISM Magnaporthe grisea

REFERENCE Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe. 1 (bases 1 to 750)

AUTHORS Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R., Phillips, K., Sabinowski, W., Wing, R.A. and Dean, R.A.

TITLE A BAC End Sequencing Framework to Sequence the Magnaporthe grisea Genome

JOURNAL Unpublished

COMMENT Contact: Dean RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson University, Clemson, SC 29634
Tel: 864 656 5737
Fax: 864 656 4293
Email: rdean@clemson.edu
Seq primer: GGAACACGCTATGACCATG
Class: BAC ends
High quality sequence stop: 344.

QY 562 AGTACCGAAGTCTA 577
 DB 85 AGCAGCGGCTCCGCCA 70

RESULT 8
 A0399120/c 583 bp DNA linear GSS 06-MAR-1999
 LOCUS mgxb0001B24f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 DEFINITION clone mgxb0001B24f, genomic survey sequence.
 ACCESSION A0399120
 VERSION A0399120.1 GI:4370147
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
 1 (bases 1 to 583)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Saslinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome

REFERENCE
 AUTHORS Unpublished
 TITLE Contact: Dean RA
 COMMENT Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdeane@clemson.edu
 Seq primer: TAAATGACTCACTATAGG
 Class: BAC ends
 High quality sequence stop: 411.
 Location/Qualifiers

FEATURES
 source 1..583
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0001B24f"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_11b="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25x genome coverage. High density colony filters
 are available upon request."

BASE COUNT 104 a 153 c 190 g 136 t
 ORIGIN

Query Match 12.3%; Score 81.4; DB 28; Length 583;
 Best Local Similarity 52.8%; Pred. No. 3.9e-08;
 Matches 200; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 154 AAAAATTCATGAAACACAAACACCAACCAAGTTGTAACATGTCATTAACATGGA 213
 DB 445 AAGAAAAAAGAGAGACTAATCAACACAAACAAAGCGCGCTCATCACTACTCG 386
 QY 214 GCCAATTCACCAACAAATGTAATGCTATTTATGGCTATGAGCTGTGACCT 273
 DB 385 GGCAGCTAAGCCCAAGGCACTATCTAGCCCTGAGGCTGAGCGCAACCCG 326
 QY 274 CTGTGCAATATATATGTCGACAGTTGGGCACTGCGCTCAACA--GGAGCAAG 330
 DB 325 CTGATGATACTACTGCTGTGAGAGCTTTGGAGCTACACCCGCTGCGGCGCAC 266

QY 331 CTAAGGGGACCATCTCTGTTGATGAGGAACATATGATATCTACGACTCTTAGATC 390
 DB 265 AACCGCGGCTCTTACCTCGAGCGGACACCTACGACATCTGTGACAGACCGCTAC 206
 QY 391 AATCAACCTCCATTTAAGGGATTGCCACATTTTAAACAAATATGAGAGTTCGAAGATCG 450
 DB 205 AACGAGCTCTCATCGACGACCAAGACCTTTTCAGCAGTTCTGTGCTGCCCGCAAC 146
 QY 451 AAACGACAGTGGCAGCATTTCTGTGACCAACCACTTTAGAGCGTGGGAAACTTAGG 510
 DB 145 AAGCGCGCAGCGGACCGTCACTTGTCCCAACGCTCAACGCTGCGGCAAGCGCGC 86
 QY 511 ATGAATATGGGAAATAT 529
 DB 85 CTCACCTCGGCAACCACT 67

RESULT 9
 A0396475/c 561 bp DNA linear GSS 06-MAR-1999
 LOCUS mgxb0010M14f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 DEFINITION clone mgxb0010M14f, genomic survey sequence.
 ACCESSION A0396475
 VERSION A0396475.1 GI:4367502
 KEYWORDS GSS.
 SOURCE Magnaporthe grisea (anamorph: Pyricularia grisea)
 ORGANISM Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
 1 (bases 1 to 561)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Saslinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome

REFERENCE
 AUTHORS Unpublished
 TITLE Contact: Dean RA
 COMMENT Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdeane@clemson.edu
 Seq primer: TAAATGACTCACTATAGG
 Class: BAC ends
 High quality sequence start: 50
 High quality sequence stop: 443.
 Location/Qualifiers

FEATURES
 source 1..561
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0010M14f"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_11b="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWICH; Site_1: HindIII; Site_2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying
 numerous aspects of the fungal-host interaction. In
 order to facilitate genome wide analysis, a BAC library
 containing 9216 clones with an average insert size of 130
 kbp was constructed. This library represents greater
 than 25x genome coverage. High density colony filters
 are available upon request."

BASE COUNT 99 a 135 c 187 g 137 t 3 others
 ORIGIN

Query Match 11.6%; Score 76.6; DB 28; Length 561;
 Best Local Similarity 53.7%; Pred. No. 4.7e-07;
 Matches 204; Conservative 0; Mismatches 170; Indels 6; Gaps 2;

QY 139 TTATTCCCTAAGGTAATAAATTCATGAAACACAAACACCAACCAAGTTGGTAACATG 198
 |||||
 Db 495 TTTTCCCAAAAGGAAAAAAGAGACTANCAAAACACCAACAAACAGCGCGT- 437
 |||||
 QY 199 TCCATAACTACGGAGCCAACTTCCAAACCAATGTAATGCGTATTTATGCGTCTATGGT 258
 |||||
 Db 436 --CATCAACTACTCGGGCAGCTACAGCCCCCGGCAACTCATCTCGCGCTTACGGC 379
 |||||
 QY 259 TGGACTGTGACCCCTTGTGCAATATATATTTGTCGACAGTTGGGGCAACTGGCGTCCA 318
 |||||
 Db 378 TGGAGCGCAACCCGCTGATCGAGTACTACGTGTGGAGAGCTTTGGCAGCTACACCCG 319
 |||||
 QY 319 CCA---GGAGNACGCCCTAAGGGACCACTACTGTGTGATGAGGAACATATGATCTAC 375
 |||||
 Db 318 TCGTGGCGCGCCACCAACCGCGGGTCTTCCCTCGACGGCAGCACCTACGACATCCTG 259
 |||||
 QY 376 GAGACTCTTAGAGTCAATCAACCTCCATTAAAGGGATGTCACATTTAAACATATTTGG 435
 |||||
 Db 258 GTCAGACCCGCTACACAGCCCTTCCATCGACGCGCACAGACCTTTTCAGAGTTCTGG 199
 |||||
 QY 436 AGTGTTCGAAGATCGAAACGACGAGTGGCGACGATTTCTGTGACCAACCACTTTAGAGCG 495
 |||||
 Db 198 TCGGTGCGCCCAACAGCGCGCAGCGCACCGTCACTTTGCGCAACCACTCAACGCC 139
 |||||
 QY 496 TGGGAAACTTAGGGATGAA 515
 |||||
 Db 138 TGGCGCAACGCGCGCTCAA 119
 |||||

RESULT 10
 AQ448084/c
 LOCUS
 DEFINITION
 clone mgxb0016B21f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 clone mgxb0016B21f, genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Magnaporthe grisea (anamorph: Pyricularia grisea)
 Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 753)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 Unpublished
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: TAATACGACTCACTATAGG
 Class: BAC ends
 High quality sequence stop: 401.
 Location/Qualifiers

FEATURES
 source
 1..753
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0016B21f"
 /tissue_type="Protoplasts"
 /lab_host="E. coli DH10B"
 /clone_lib="CUGI Rice Blast BAC Library"
 /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
 Rice blast is one of the most devastating fungal diseases
 of rice world wide. It is a filamentous ascomycete with
 a haploid genome (n=7) of approximately 40 Mbp. Rice
 blast is an important model fungal pathogen for studying

BASE COUNT 140 a 202 c 253 g 158 t
 ORIGIN
 Query Match 11.3%; Score 74.6; DB 28; Length 753;
 Best Local Similarity 55.1%; Pred. No. 1.4e-06;
 Matches 168; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

QY 201 CATAACTACGGAGCCAACTTCCAAACCAATGTAATGCGTATTTATGCGTCTATGGTTG 260
 |||||
 Db 396 CATCACTACTCGGGCAGCTACAGCCCCCGGCAACTCATCTGGCGGTCTACGGGTG 337
 |||||
 QY 261 GACTGTGTGACCCCTTCTCGAATATATATTTGTCGACAGTTGGGGCAACTGGCGTCCACC 320
 |||||
 Db 336 GACGCGCAACCCGCTGATCGAGTACTAGTGTGGAGAGCTTTGGCAGCTACAACCCGTC 277
 |||||
 QY 321 A---GGAGCAACGCTTAAGGGGACCATCACTGTTGATGAGGAAACATATGATATCTACGA 377
 |||||
 Db 276 GTCGGGCGCCACCATCCGCGGCTCTTCACTCGGACGGCAGCACCTACGACATCCTGT 217
 |||||
 QY 378 GACTCTTAGAGTCAATCAACCCCTCCATTAAGGGGATTTGCCACATTTAAACAATATTGGAG 437
 |||||
 Db 216 CAGCACCGCTTACAAACGAGCCCTCATCGACGGCACCAGACCTTTTCAGCAGTTCTGTC 157
 |||||
 QY 438 TGTTCGAAGATCGAAACGACGAGTGGCAGCAATTTCTGTCAGCAACCACTTTAGAGCGTG 497
 |||||
 Db 156 GGTGCGCGCAACAGCGCGCCAGCGCACCTTTGCCAACCAACGCTCAACGCGCTG 97
 |||||
 QY 498 GGAAG 502
 |||||
 Db 96 GCACA 92
 |||||

RESULT 11
 AQ398756/c
 LOCUS
 DEFINITION
 clone mgxb0005L16f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
 clone mgxb0005L16f, genomic survey sequence.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Magnaporthe grisea (anamorph: Pyricularia grisea)
 Magnaporthe grisea
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 1 (bases 1 to 418)
 Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
 Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
 A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
 Genome
 Unpublished
 Contact: Dean RA
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson University, Clemson, SC 29634
 Tel: 864 656 5737
 Fax: 864 656 4293
 Email: rdean@clemson.edu
 Seq primer: TAATACGACTCACTATAGG
 Class: BAC ends
 High quality sequence stop: 334.
 Location/Qualifiers

FEATURES
 source
 1..418
 /organism="Magnaporthe grisea"
 /mol_type="genomic DNA"
 /strain="70-15"
 /db_xref="taxon:148305"
 /clone="mgxb0005L16f"


```

REFERENCE 1 (bases 1 to 561)
AUTHORS Zhang,H., Potokina,E., Michalek,W., Weschke,W., Stein,N. and Graner
        A.
TITLE Barley ESTs from germinating seeds
JOURNAL Unpublished
COMMENT Contact: Stein Nils
        Molecular Markers Group, Department Genbank
        Institute of Plant Genetics and Crop Plant Research (IPK)
        Corrensstr. 3, 06466, Gatersleben, Germany
        Tel: 039482-5522
        Fax: 039482-5595
        Email: stein@ipk-gatersleben.de
        Insert Length: 561 Std Error: 0.00
        Place: 4 row: A column: 2
        Seq primer: M13rev.

FEATURES             Location/Qualifiers
     source           1..561
                     /organism="Hordeum vulgare subsp. vulgare"
                     /mol_type="mRNA"
                     /cultivar="barke"
                     /db_xref="taxon:112509"
                     /clone="HV04A02"
                     /tissue_type="germinating seeds"
                     /dev_stage="germinating seeds (48-96 h)"
                     /lab_host="XL10-Gold"
                     /clone_lib="HV"
                     /note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of
cDNA); Site 2: XhoI (3'-end of cDNA); Roots were grown for
two days on filter paper at room temperature. Due to a
cloning artefact caused by the kit, in most cases the
EcoRI site is NOT present, as well as the EcoRI adapter
used for cloning. To excise the insert, restriction sites
upstream EcoRI should be used (e.g. BamHI, SalI, PstI).
NOTE: Also due to the cloning system used Blue/white
selection for recombinants is not 100% reliable. Average
insert size is 1 kb"
BASE COUNT  117 a  184 c  148 g  112 t
ORIGIN
Query Match      10.0%; Score 66.2; DB 13; Length 561;
Best Local Similarity 57.7%; Pred. No. 0.00011;
Matches 138; Conservative 0; Mismatches 98; Indels 3; Gaps 1;

QY 200 CCATAACTACGAGCCAACTCCACCAATGGTAATCGGTATTATTCGGTCTATGGTT 259
Db 320 CGATCACTACGGCGGATCTTTCCAGCCTCAGGGTAACGGTTACCTTCCGGTCTACGGTT 379

QY 260 GGACTGTTGACCTCTTTCGNAATATTATATGTCGACAGTTGGGGCACTGGCGTCCAC 319
Db 380 GGACCCGCAACCCGCTTGTGGAGTACTACGTGATCGAGTCTTACGGCACCTTACAACCCCA 439

QY 320 CA---GGAGCAACGCTTAAGGGGACCATCACTGTTGATGGAGGACATATGATATCTACG 376
Db 440 GCAGCGGATCCAGCGGAGGCGAGCTTCAAGCTGACGGCGGTACCTACGACGTTTCCA 499

QY 377 AGACTCTTAGAGTCAATCAACCTCCATTAAAGGGGATTGCCACATTTAAACAATATGG 435
Db 500 CCTCACCCTGACCAACCAAGCCCTCCATCGATGGAACAAGGACCTTTCAGCAGTACTGG 558

RESULT 14
CD464005 473 bp mRNA linear EST 04-JUN-2003
LOCUS ETH1_48_B06.bl_A002 Ethylene-treated seedlings Sorghum bicolor cDNA
DEFINITION clone ETH1_48_B06_A002 3', mRNA sequence.
ACCESSION CD464005
VERSION CD464005.1 GI:31385273
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.

```

```

REFERENCE 1 (bases 1 to 473)
AUTHORS Cordonnier-Pratt,M.-M., Wentzel,V., Suzuki,Y., Sugano,S., Klein
        R.R., Liang,C., Sun,F., Sullivan,R., Shah,M., Salzman,R., Chua Tan
        N., Gonzalez,M., Lane,S., Miller,V., Nanda,P., Olaseinde,O.,
        Eastman,A. and Pratt,L.H.
TITLE An EST database from Sorghum: 1-aminocyclopropane-1-carboxylic acid
        (ACC)-treated seedlings
JOURNAL Unpublished
COMMENT Other ESTs: ETH1_48_B06.g1_A002
        Contact: Cordonnier-Pratt MM
        Laboratory for Genomics and Bioinformatics
        The University of Georgia, Department of Plant Biology
        Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
        Tel: 706 542 1860
        Fax: 706 583 0210
        Email: mmpratt@uga.edu
        Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
        the Human Genome Center, University of Tokyo Institute of Medical
        Science; plant material and RNA prepared at Texas A & M University;
        sequencing done in the Laboratory for Genomics and Bioinformatics,
        University of Georgia. Sequence ends have been trimmed to exclude
        vector and regions below Phred quality 16. Three-prime sequences
        are presented as their reverse complement and have been trimmed to
        exclude polyA.
        Seq primer: Sug3 (CGACCTGCAGCTCGAGCACA)
        POLYA=Yes.

FEATURES             Location/Qualifiers
     source           1..473
                     /organism="Sorghum bicolor"
                     /mol_type="mRNA"
                     /cultivar="BTX623"
                     /db_xref="taxon:4558"
                     /clones="ETH1_48_B06_A002"
                     /lab_host="DH10B-T1 phage-resistant E. coli"
                     /clone_lib="Ethylene-treated seedlings"
                     /note="Vector: pME18S-FL3; Site 1: XhoI; Site 2: XhoI; The
library was prepared from polyA+ RNA from seedlings grown
in hydroponic culture. At 8 days of age, medium was
supplemented with 0.5 mM 1-aminocyclopropane-1-carboxylic
acid (ACC) to induce endogenous ethylene (ETH) production.
Roots and shoots were harvested after 27 and 72 hr and
material from both time points was combined prior to RNA
isolation. Double-stranded cDNA was cloned
unidirectionally into different DraIII sites of the
pME18S-FL3 vector (5-prime DraIII site is CACTGTGTG,
3-prime DraIII site is CACCATGTG). XhoI excises the cDNA
insert."
BASE COUNT  109 a  126 c  112 g  126 t
ORIGIN
Query Match      9.5%; Score 63.2; DB 14; Length 473;
Best Local Similarity 53.7%; Pred. No. 0.0005;
Matches 131; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 334 AAGGGGACCATCACTGTTGATGGAGGAACATATGATATCTACGAGACTCTTTAGAGTCAAT 393
Db 16 AAGGGTGAGGTCACTGTTGACGATCTGTTTACGACATTTACGTACGACCCCGTGTCAAC 75

QY 394 CAACCTCCATTAGGGGATGCCACATTTAAACAATATTGGAGTGTTCGAAGATCGAAA 453
Db 76 GCCCCTCCATTAGGGGTAACAAGCCCTTTTGGTCTCTTCGACGCAACAAG 135

QY 454 CGCACGAGTGGCAGCATTTCTGTGCAGCAACACACTTTAGAGCGTGGAAAACCTTAGGGATG 513
Db 136 CGATCCAGCGGATCCGTCACACCGGGTCTCATTTCAGGCGCTGGAAAATATGTGGCGCTC 195

QY 514 AATATGGGAAAATGTATGAAGTCGCGCTTACTCTAGAGAGGCTATCAAGTAGCGGAAGT 573
Db 196 ACCCTTGTTACCAACAACATATCAGATCCTTCTGTTCGGGGCTACTACAGCTCTGGCTCT 255

QY 574 GCTA 577
Db 256 GCCA 259

```

RESULT 15
 BO664593/c
 LOCUS
 DEFINITION HV04A02u HV Hordeum vulgare subsp. vulgare cDNA clone HV04A02
 3-PRIME, mRNA sequence.
 ACCESSION BO664593
 VERSION BO664593.1 GI:21806275
 KEYWORDS EST.
 SOURCE Hordeum vulgare subsp. vulgare
 ORGANISM Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Scrophophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.
 1 (bases 1 to 483)
 Zhang, H., Potokina, E., Michalek, W., Weschke, W., Stein, N. and Graner
 A.
 TITLE Barley ESTs from germinating seeds
 JOURNAL Unpublished
 COMMENT Contact: Stein Nils
 Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595
 Email: stein@ipk-gatersleben.de
 Insert Length: 483 Std Error: 0.00
 Plate: 4 row: A column: 2
 Seq primer: M3uni.
 Location/Qualifiers
 1.483
 /organism="Hordeum vulgare subsp. vulgare"
 /mol_type="mRNA"
 /cultivar="Baiké"
 /db_xref="GABI:148372"
 /db_xref="taxon:112509"
 /clone="HV04A02"
 /issue_type="germinating seeds"
 /dev_stage="germinating seeds (48-96 h)"
 /lab_host="XLI0-Gold"
 /clone_1lb="HV"
 /note="Vector: pBluescript SK+; Site_1: EcoRI (5'-end of
 cDNA); Site_2: XhoI (3'-end of cDNA); Roots were grown for
 two days on filter paper at room temperature. Due to a
 cloning artefact caused by the kit, in most cases the
 EcoRI site is NOT present, as well as the EcoRI adapter
 used for cloning. To excise the insert, restriction sites
 upstream EcoRI should be used (e.g. BamHI, SalI, PstI).
 NOTE: Also due to the cloning system used Blue/white
 selection for recombinants is not 100% reliable. Average
 insert size is 1 kb"

BASE COUNT 102 a 130 c 140 g 111 t
 ORIGIN

Query Match 9.1%; Score 60.4; DB 13; Length 483;
 Best Local Similarity 51.2%; Pred. No. 0.0021;
 Matches 167; Conservative 0; Mismatches 156; Indels 3; Gaps 1;

271 CCTCTTGTGAATATATATATGTCAGATGGGCACTGGCGTCC--ACGAGAGCA 327
 DB 483 CCGCTTGTGAATATATATGTCAGATGGGCACTGGCGTCC--ACGAGAGCA 424
 328 ACGCCTAAGGGGACATCACTGTGATGAGAGAACATATATCTACGAGACTTTAGA 387
 DB 423 CAGCGGAAGGGGACATCACTGTGATGAGAGAACATATCTACGAGACTTTAGA 364
 388 GTCAATCAACCTCTCAATTAAGGGGATGCCAATTAACATATTTGAGAGTTGCAAGA 447
 DB 363 ACCAACCACCCCTCAATTAAGGGGATGCCAATTAACATATTTGAGAGTTGCAAGA 304
 448 TCGAAGCGACAGATGTCAGATTTCTGTAGACACCACTTTAGAGCGTGGGAAACTTA 507

DB 303 CAGAACCCGTCGAGGACGCTCAACATGCAAAAACCACTTCACGCTGGGCTGTTAC 244
 QY 508 GGGATGAATATGGGAAATGTATGAGTCCGCTTACTGTAGAGGCTATCAAGTAGC 567
 DB 243 GGCATGAATCTGGGCGACACTACTACAGATTGTGCGCAGTGAAGGTTACAGTCTTCT 184
 QY 568 GGAAGTCTAATGTATATAGCAATTC 593
 DB 183 GGCACCTCGACATCTATGTGCAGAC 158

Search completed: October 20, 2003, 12:03:43
 Job time : 2051 secs